



# An Idea to Address Uncertainty in Mapping at the RI Stage of the LPR Project



Meeting of CPG and EPA Region 2  
January 27, 2016

# Outline

- Introduction
- Overview of applied geostatistical approaches
- Application of conditional simulation
  - Illustrated with preliminary results

# Maps of Sediment COC Concentrations are Basis for Crafting & Evaluating Remedial Alternatives

- Supported by an array of data
  - Contaminant concentrations
  - Sediment type
  - Bathymetry
  - Long-term erosion/deposition patterns

# Maps Only Provide Estimates of the True Concentration Patterns

- On average have 0.5 samples per acre of river bottom
- Estimates at unsampled locations can have considerable error (uncertainty)

# Uncertainty is Acceptable for FS

- Recognized and accepted fact at the FS stage of a CERCLA project
- Constrained by knowledge of the river
- Favorable test of map at RM 10.9

**Surface-weighted Area Concentration  
Estimates for RM 10.9 Design Area**

	<b>Exclude Design Data</b>	<b>Include All Data</b>
Cores per acre	0.9	7.7
Pre-remedial SWAC (ng/kg)	3,361	3,179
Post-remedial SWAC (ng/kg)	85	95
Percentage SWAC reduction	97%	97%
Target area (acres)	6.1	5.1
Non-target area (acres)	6.9	7.9
Net percent area change	—	7.3%
SWAC outside footprint (ng/kg)	159	157
SWAC within footprint (ng/kg)	7,022	7,835

# But, CPG Recognizes Other Region 2 Concerns With CPG Thiessen Polygon Maps

- Magnitude of uncertainty outside of RM 10.9
- Possibility for high bias in estimates of remedy effectiveness
  - Overstating magnitude of high concentrations
  - Understating magnitude of low concentrations

# To Quantify Uncertainty and Address Potential Bias, CPG Has Explored the Following

- Conditional simulation based on kriging

# Inspired by R2 White Paper & Approaches Used Elsewhere

- R2 WP simulation illustrating uncertainty and bias issues
- Oil & gas and mining industries mapping of deposits
- EPA recommended method for characterizing wastes (EPA/600/R-92-033)
- EPA approach to target sampling to reduce uncertainty at East Poplar Creek & Lower Fox River
- EPA estimate of uncertainty of contaminated sediment volume at Trenton Channel
- EPA crafting of remedial alternatives at Kalamazoo River
- GE and EPA evaluating exposure concentrations for the Hudson River floodplain
- EPA explored adequacy of upper bound estimates of mean concentration in the Lower Duwamish Waterway

# Proposed Uses of Conditional Simulation

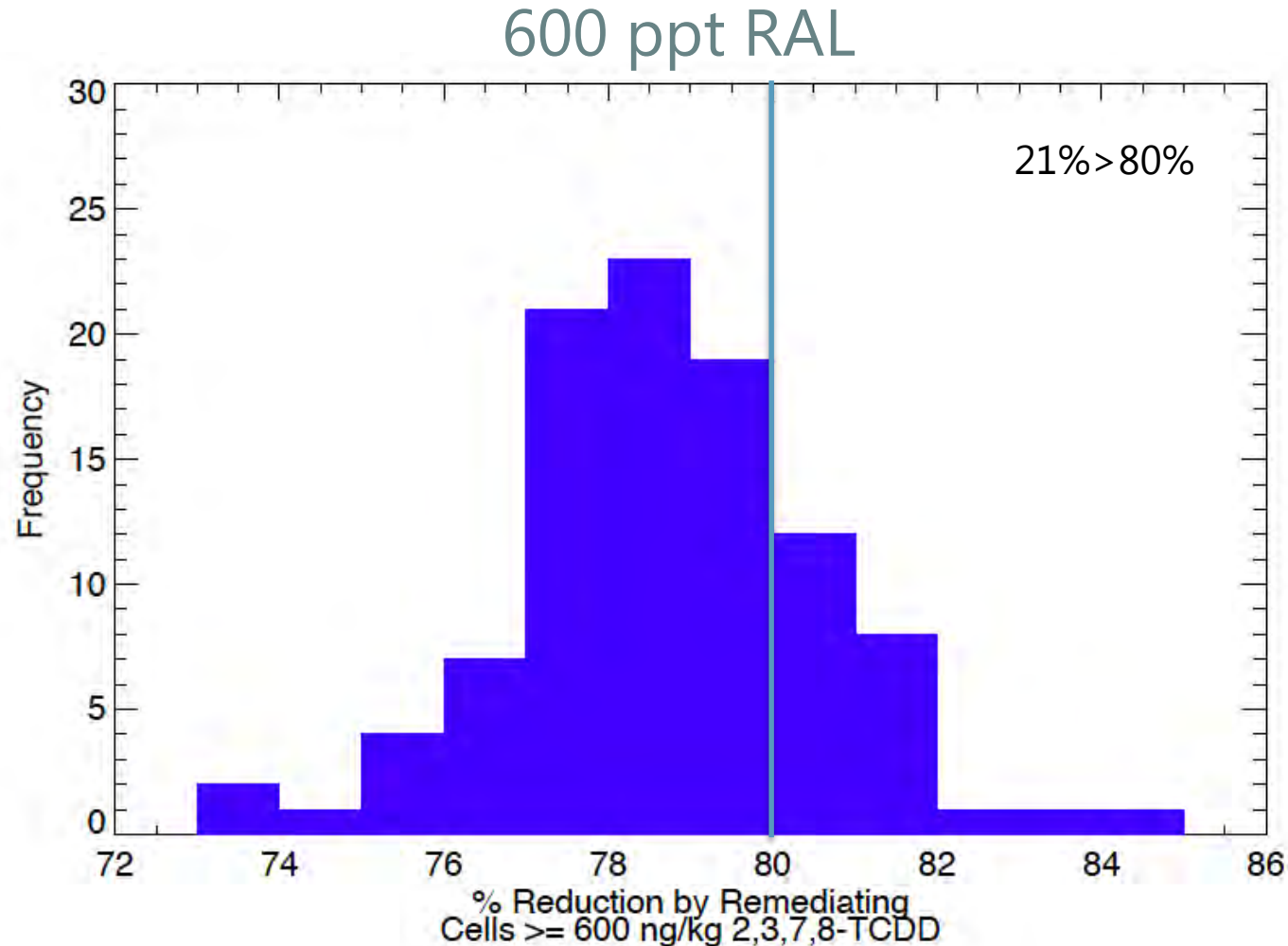
- Develop 100 plausible maps of concentrations
- Use maps to support crafting remedial options
  - Based on the 100 estimates of concentration reduction associated with any remedial action level (RAL)
- Use maps to inform data collection during remedial design
  - Identify areas with greatest uncertainty relative to RAL and target with greatest sampling density

# Use of Conditional Simulation to Craft Remedial Options for FS Evaluation

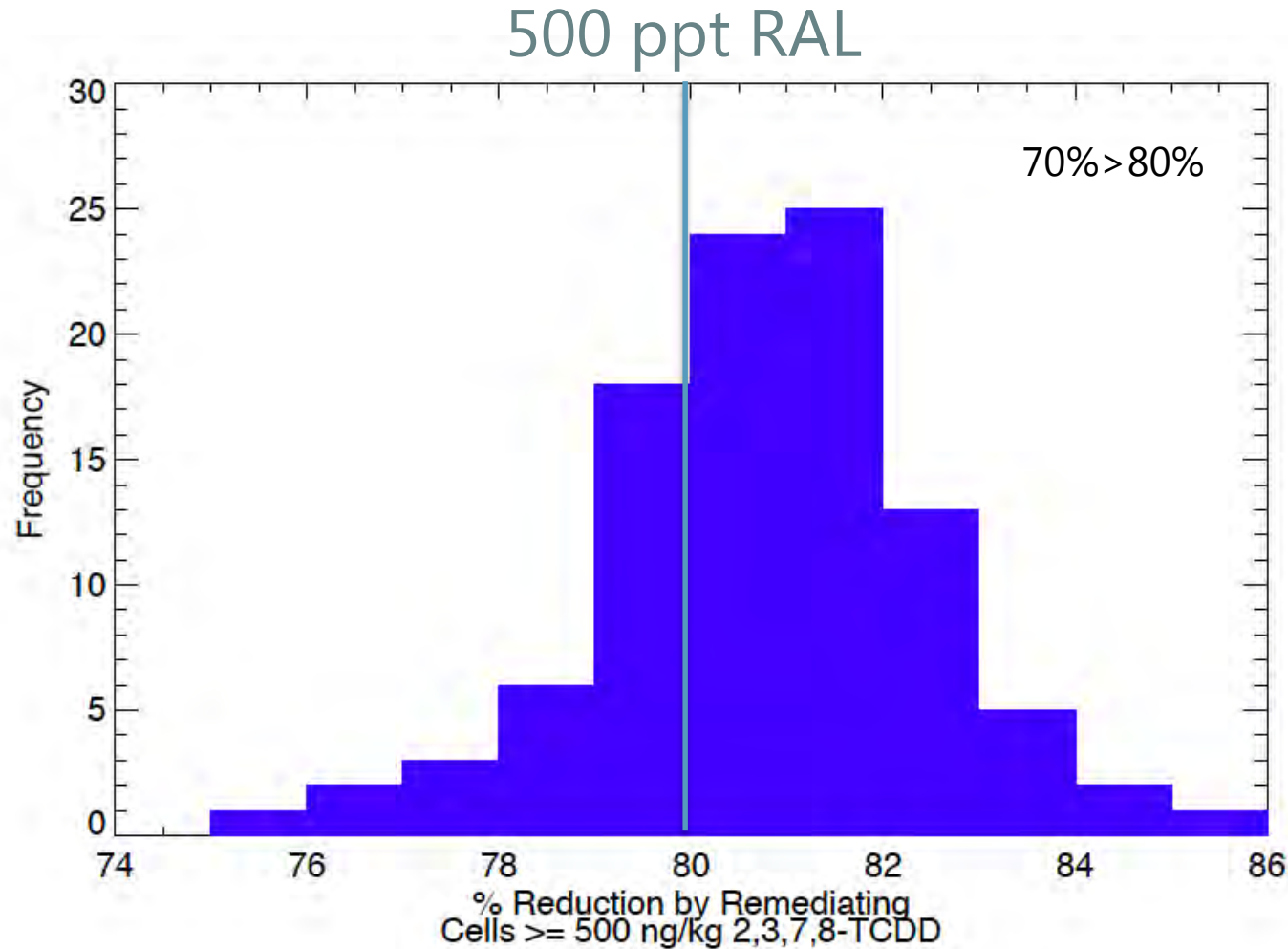
- Choosing an RAL
  - Could choose RAL that achieves greater than a specified reduction with a define level of confidence (e.g., 80% chance of achieving more than an 80% reduction)
- Choosing an area to target at a given RAL
  - Could choose conservative estimate of area meeting an RAL (e.g., 80% upper bound on area)

*Results that follow to illustrate these ideas are based on CPG initial efforts that are subject to refinement*

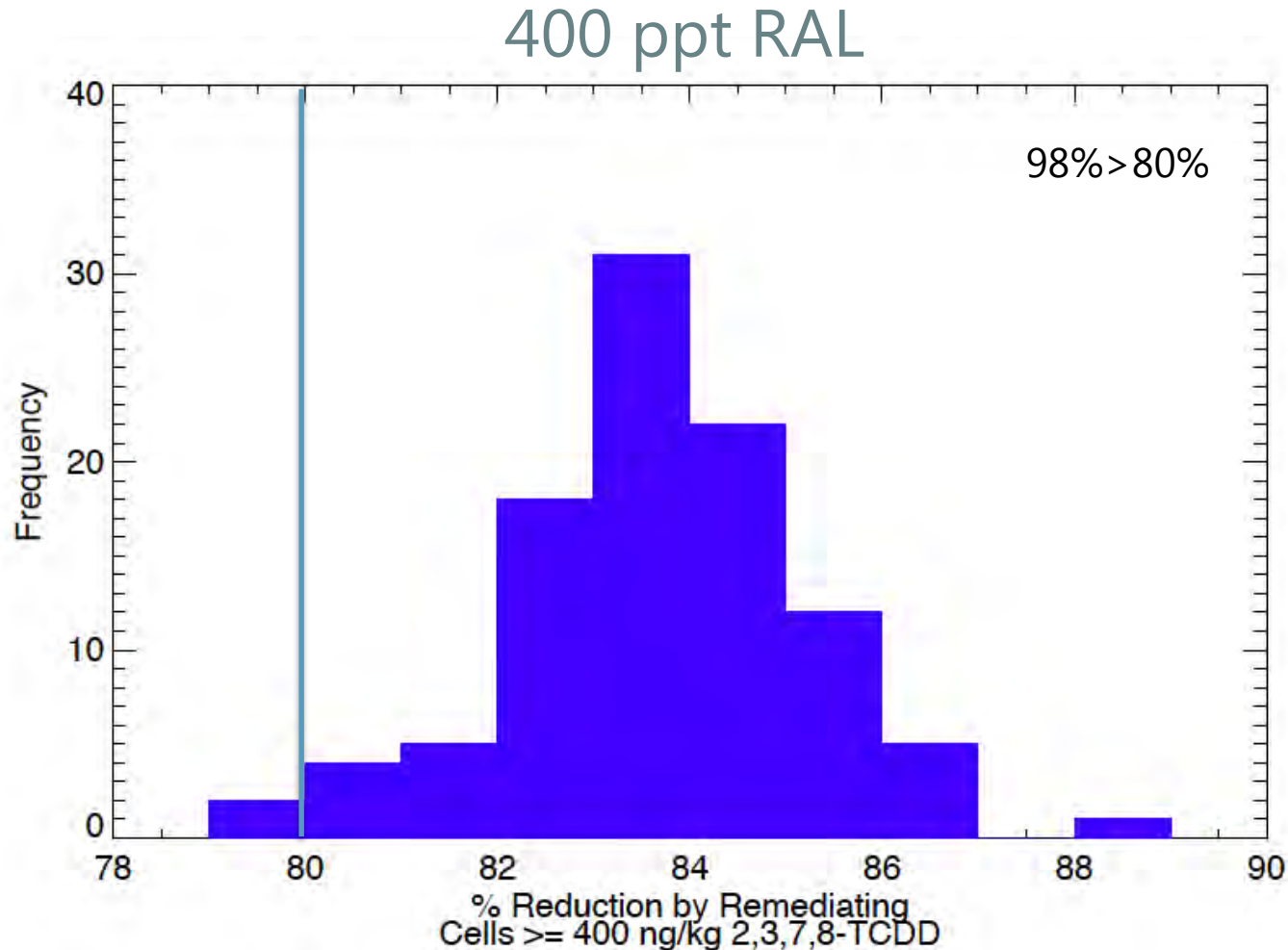
# Range of Possible Concentration Reductions Give Perspective on Uncertainty



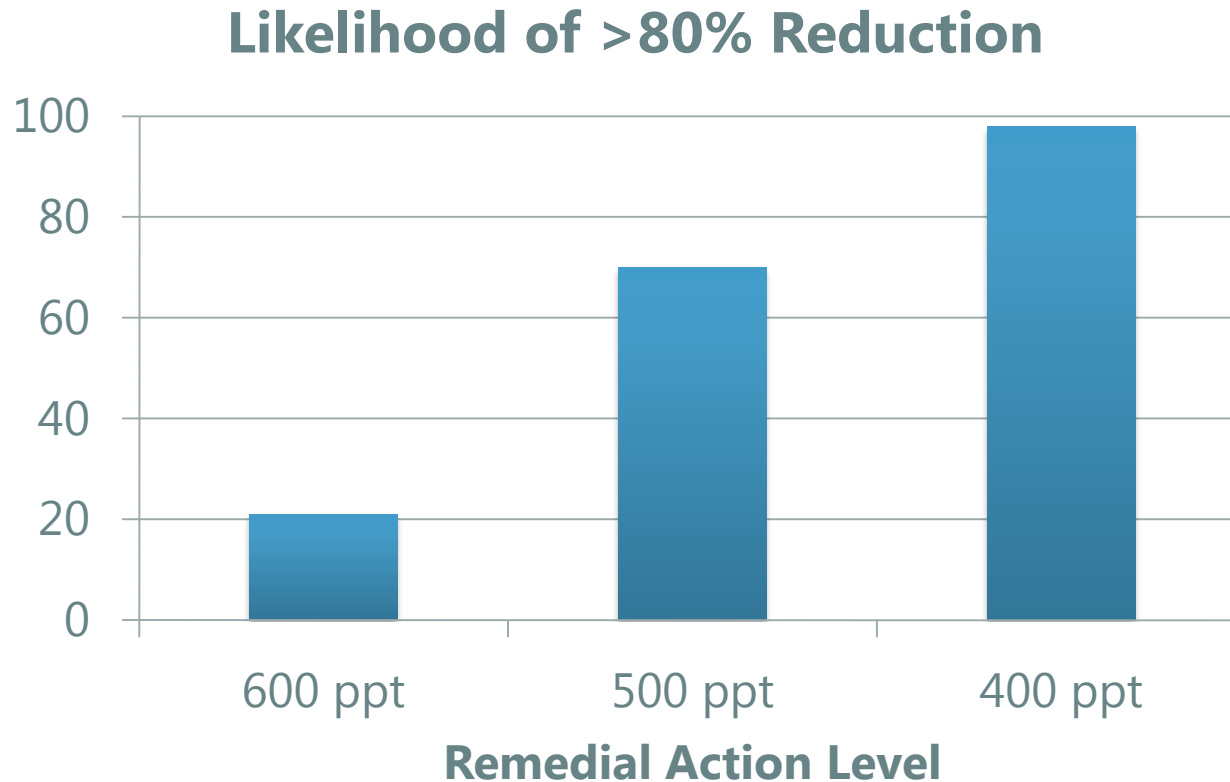
# Range of Possible Concentration Reductions Give Perspective on Uncertainty



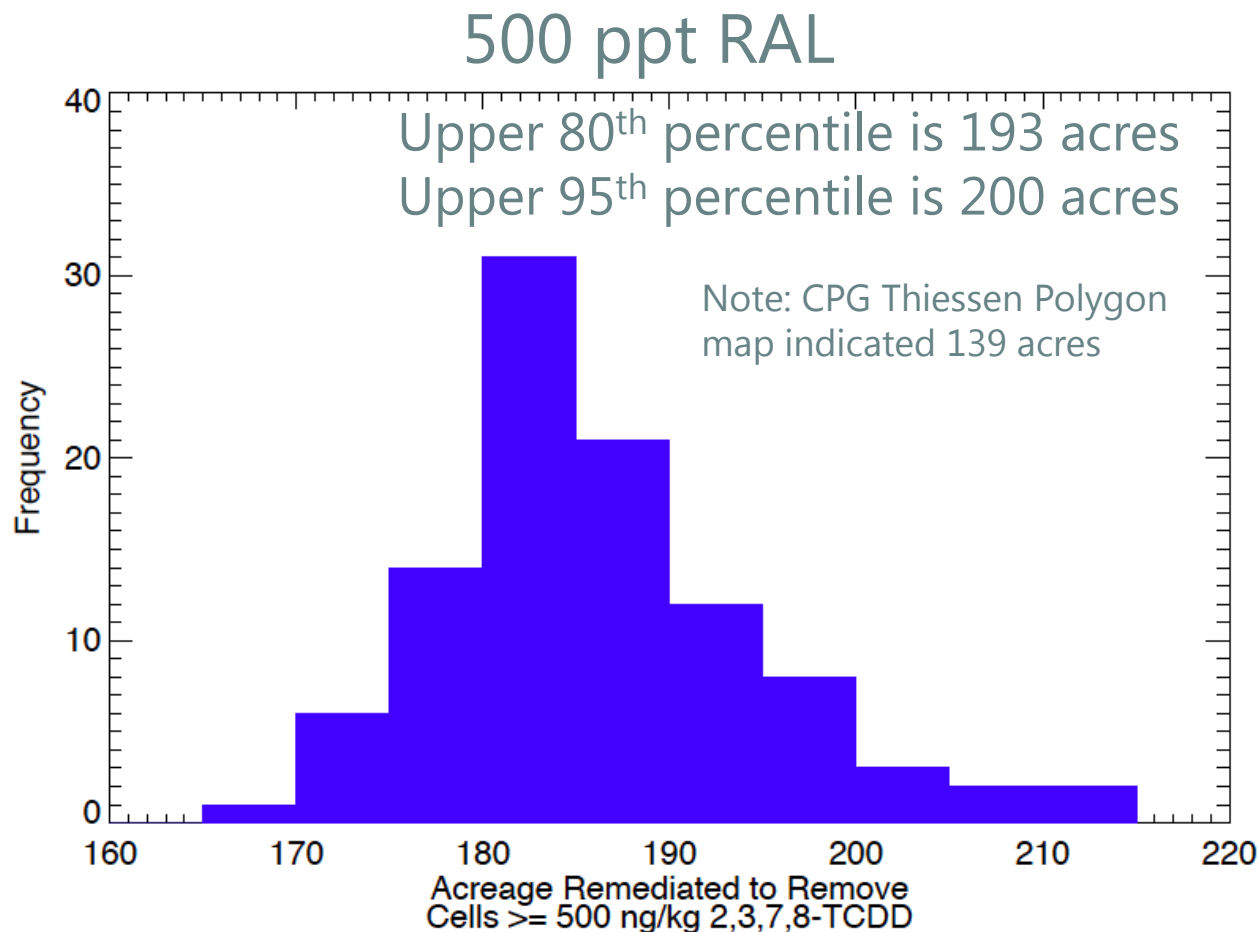
# Range of Possible Concentration Reductions Give Perspective on Uncertainty



# Range of Possible Concentration Reductions Give Perspective on Uncertainty



# Uncertainty in Area Meeting an RAL Informs Choice of Area to Characterize a Remedial Option

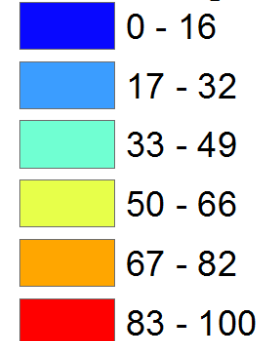


CS results can provide basis to focus design sampling

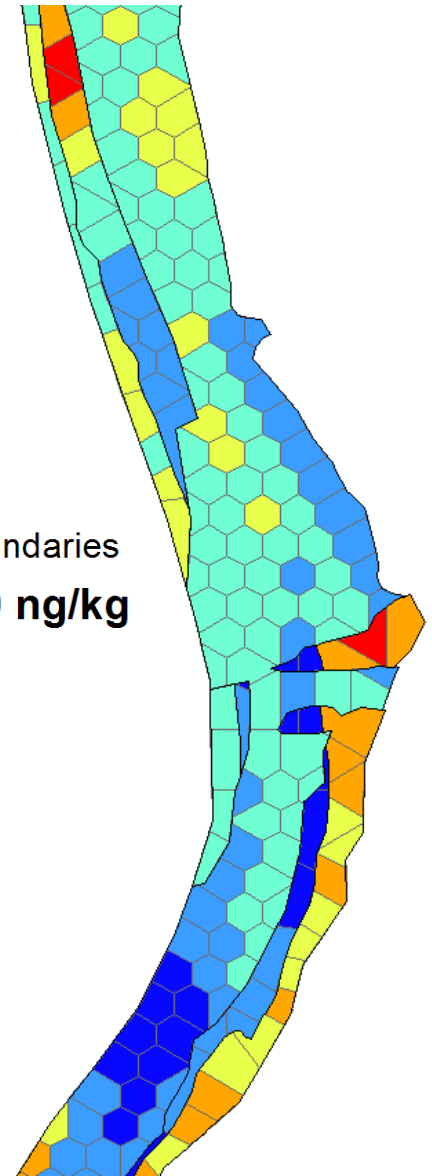
Greatest density in areas with greatest uncertainty about meeting an RAL (e.g., 33 to 66 percent chance – cyan & yellow in the figure)

Interpolation Boundaries

**Percentage  $\geq$  500 ng/kg**



RM 9.5



# Outline

- Introduction
- Overview of applied geostatistical approaches
- Application of conditional simulation
  - Illustrated with preliminary results

# Overview of Applied Geostatistical Approaches

- Kriging

- Interpolate on fine grid using measured values and a model of spatial correlation (variogram)
- Predict a distribution of possible concentrations at each grid location
- “kriging estimates present a serious drawback well known by geostatisticians as the smoothing effect in which small values are usually overestimated and large values underestimated... .. As a consequence of the smoothing effect ordinary kriging estimates do not reproduce either the histogram or the spatial variability as given by the semivariogram function.” - Yamamoto, 2005
- Kriged means/medians are not realistic concentration fields and should not be used to assess a Targeted Remedy

# Overview of Applied Geostatistical Approaches

- Conditional simulation
  - Uses kriging distributions and the observed data to create random concentration fields
  - These fields reproduce the data distribution and spatial variability as defined by the semivariogram function; they are realistic concentration fields
  - Each random field is equally probable

# Outline

- Introduction
- Overview of applied geostatistical approaches
- Application of conditional simulation
  - Illustrated with preliminary results

# Steps in Implementing Conditional Simulation

- I. Segment the River
- II. Develop variograms
- III. Krige
- IV. Conditional Simulation
- V. Interpret Results

# Segment the River

- Account for major features
  - Shoal and channel
  - Geomorphic features
- Try to preserve stationarity of concentration field (fixed mean)

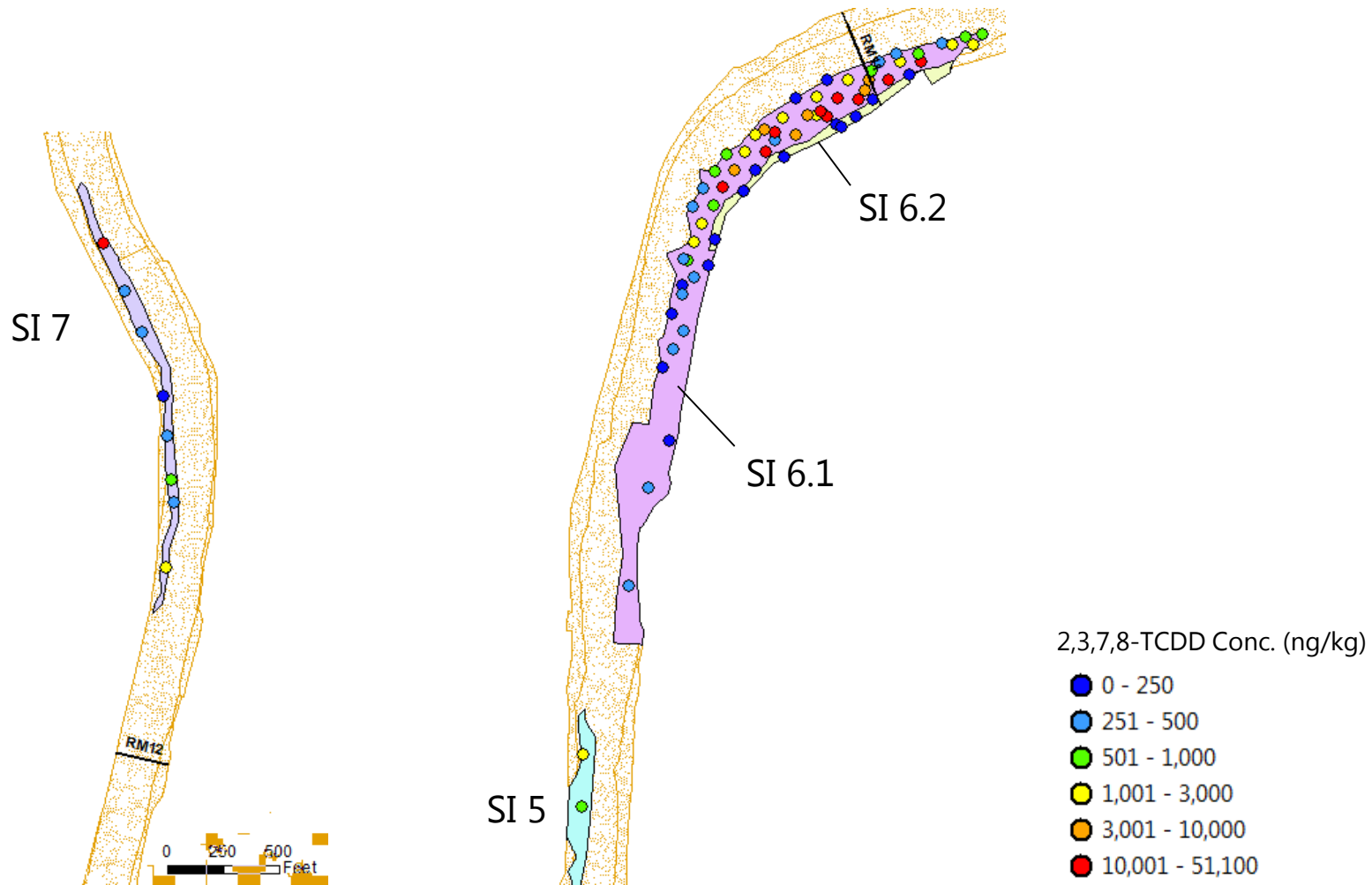
# River Segmentation – Upstream of RM 7.8

- Silt
  - Split into individual silt deposits
- Shoal/Channel
  - Split at gaps (i.e., where silt crosses the shoal/channel)
  - Split at EPA geomorphic breaks
  - Split at concentration pattern breaks

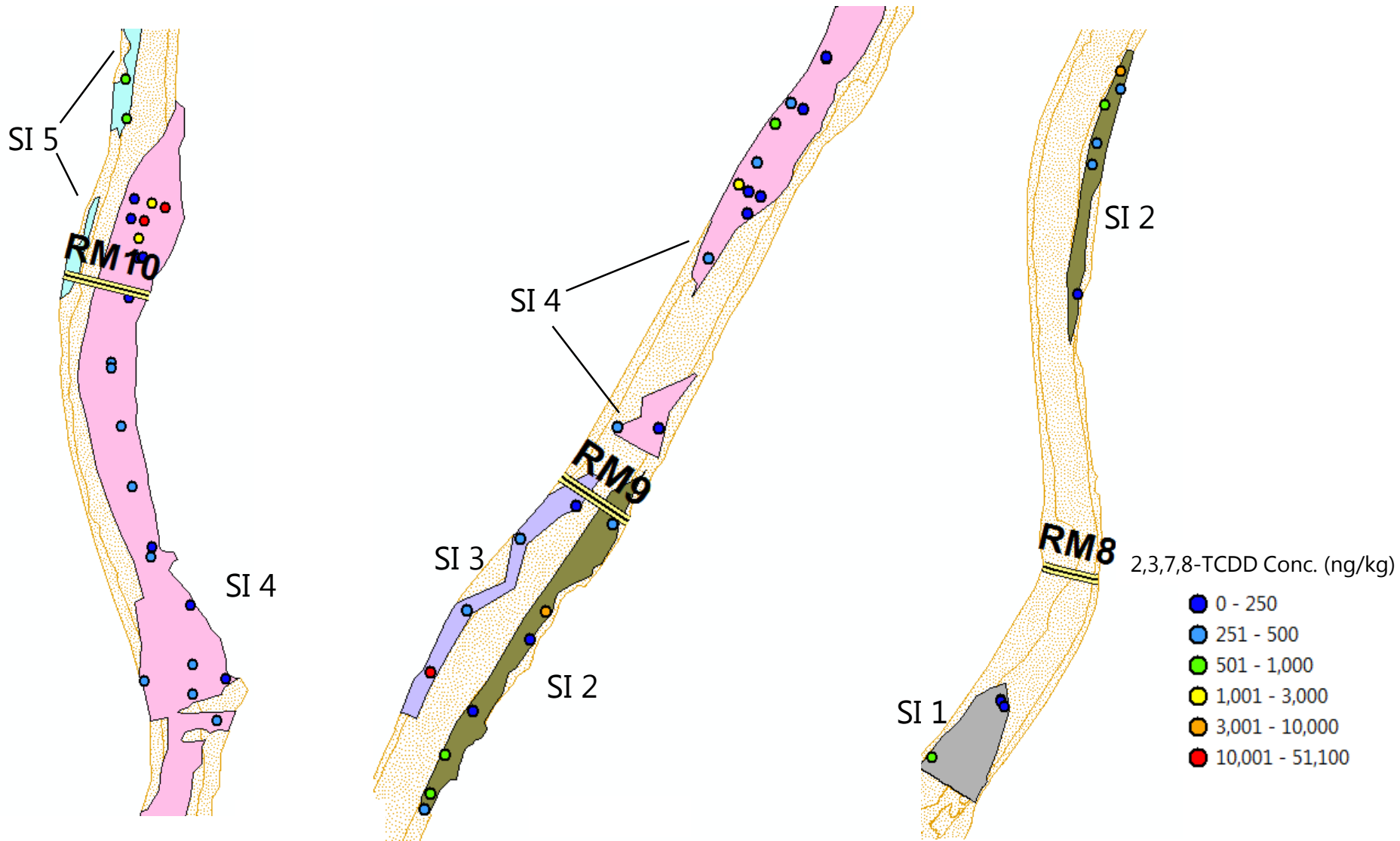
# River Segmentation – Downstream of RM 7.8

- Shoal
  - Split at EPA geomorphic breaks
- Channel groupings
  - Bathymetry-based (RM 2.3-7.8)
  - Channel downstream of RM 2.3
  - No additional subdivisions within these groups

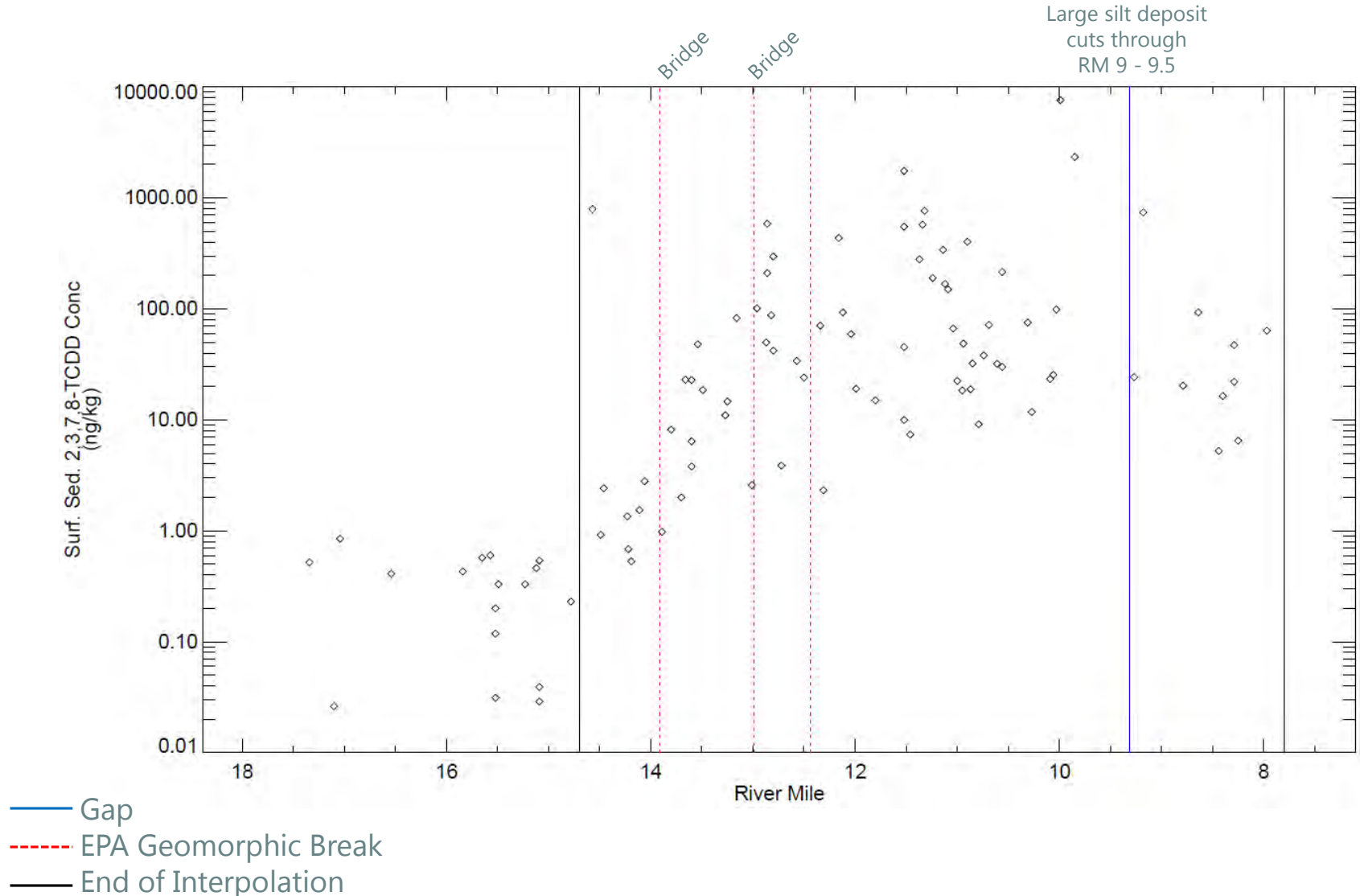
# River Segmentation – Silt Upstream of RM 7.8



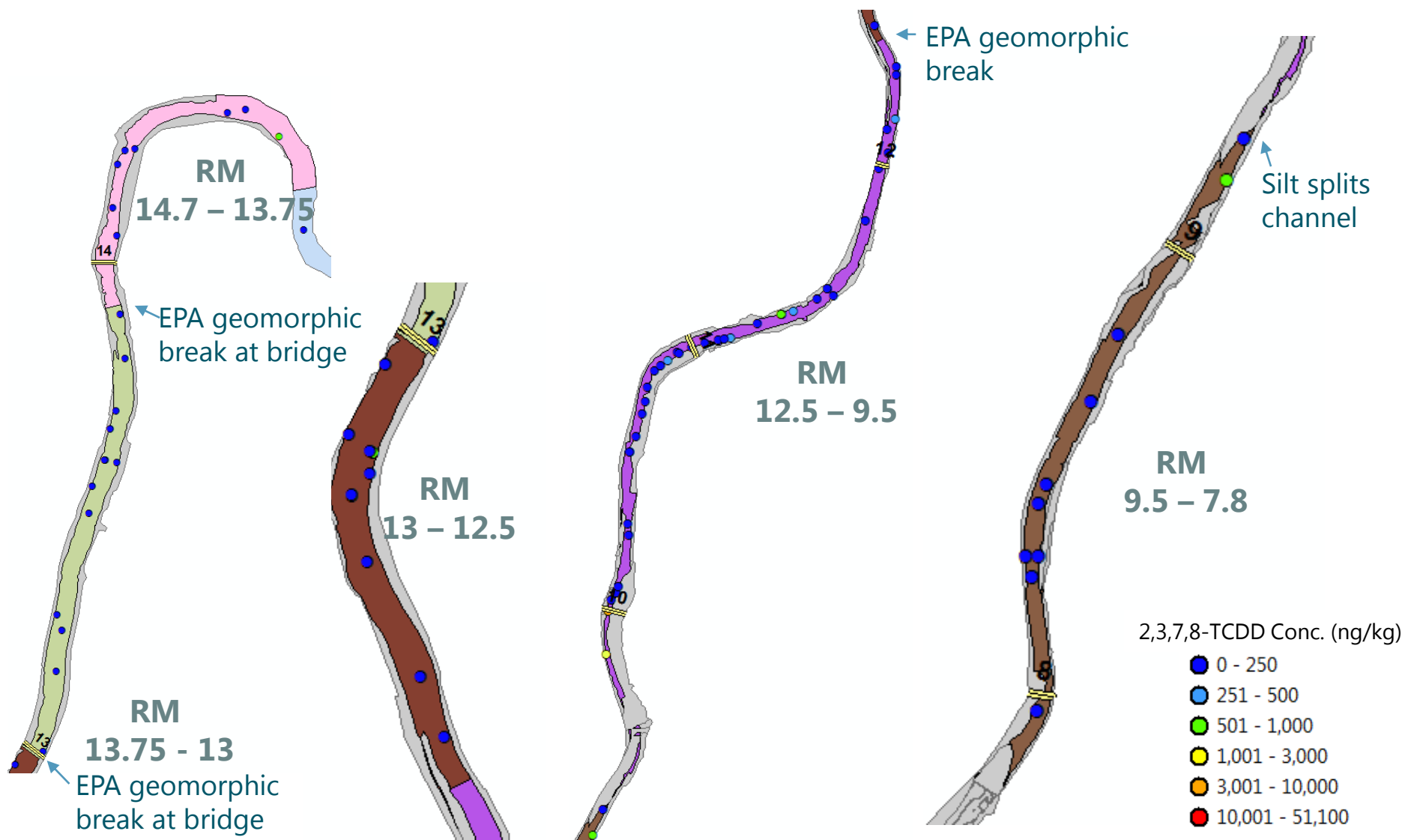
# River Segmentation – Silt Upstream of RM 7.8



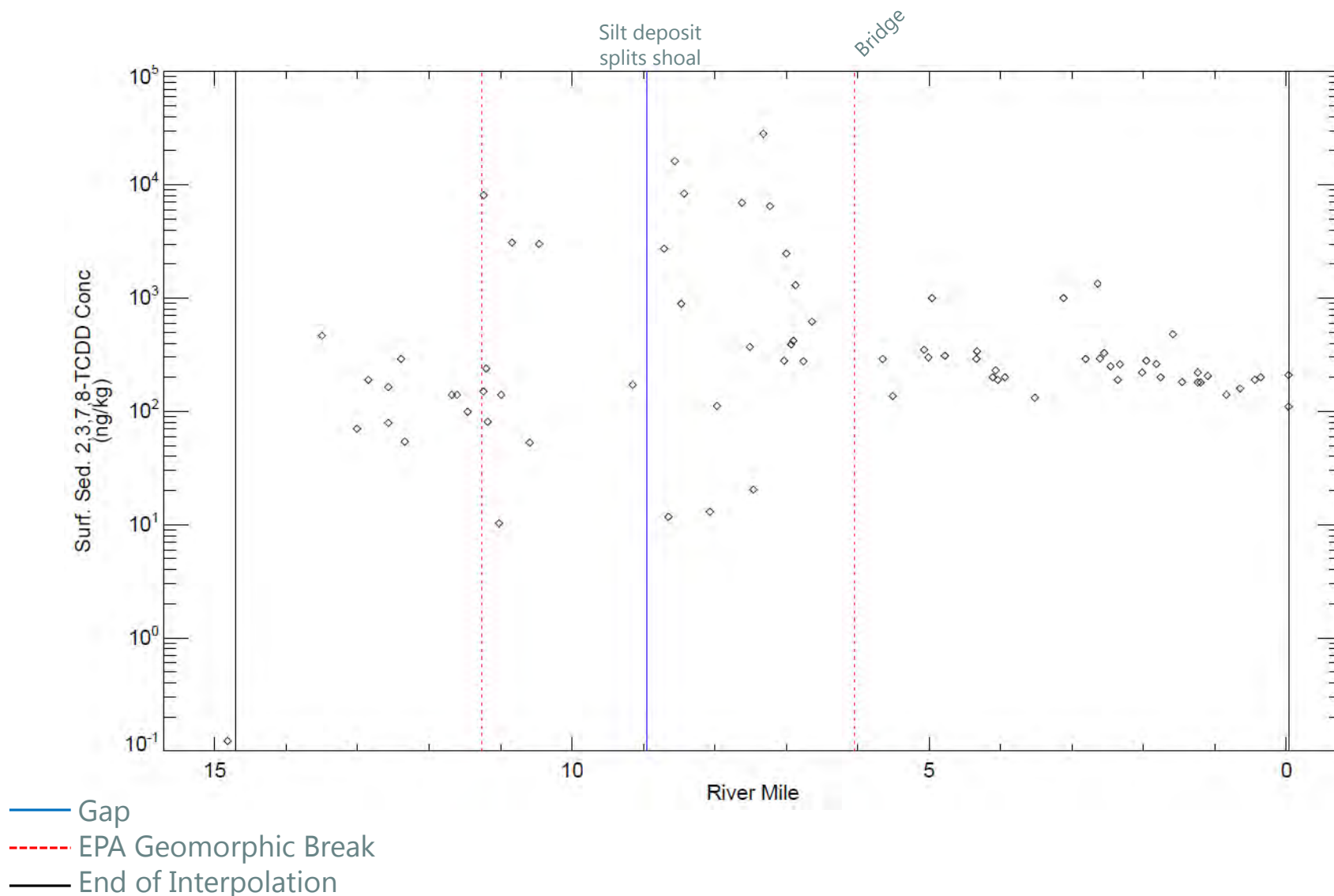
# River Segmentation – Channel Upstream of RM 7.8



# River Segmentation – Channel Upstream of RM 7.8



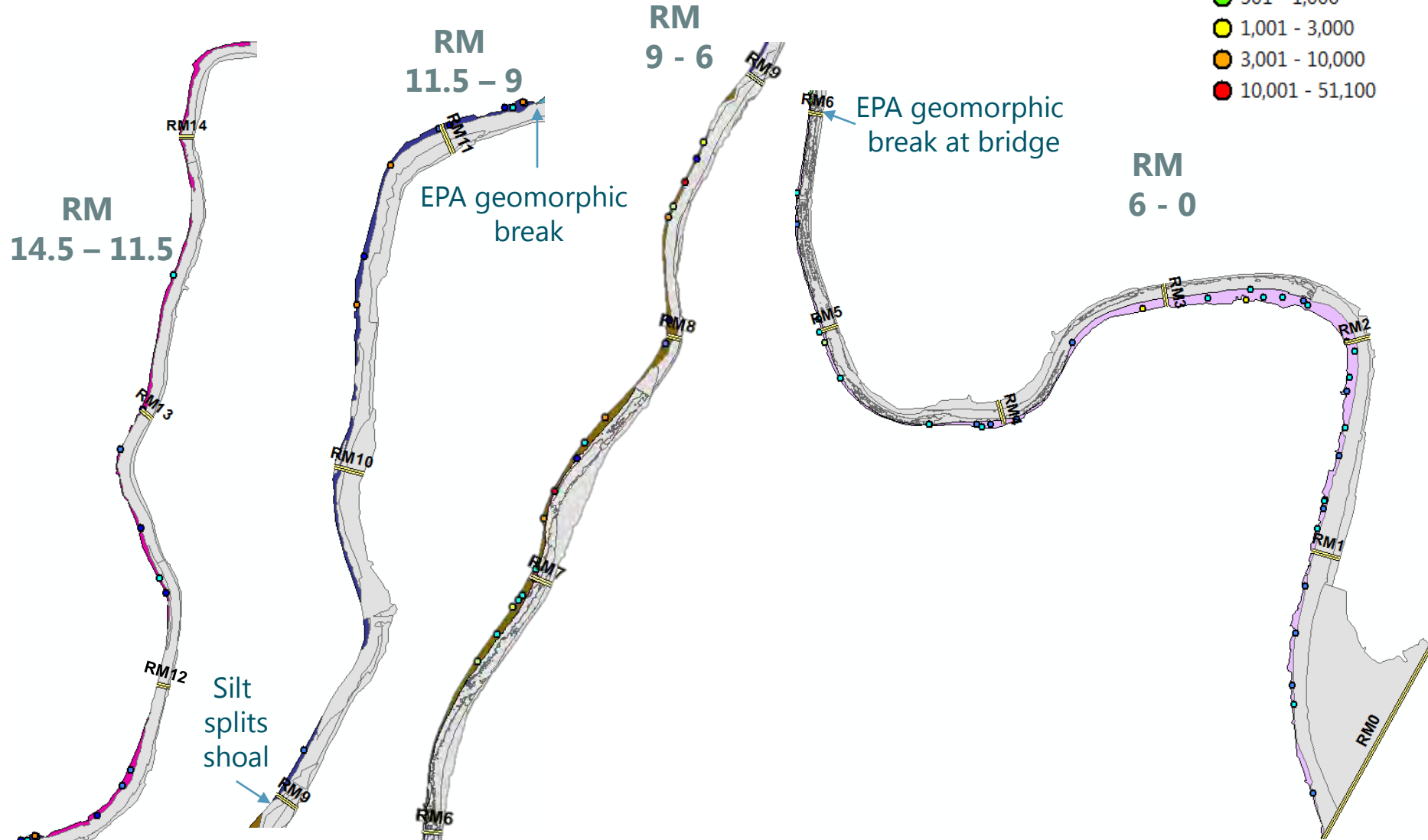
# River Segmentation – Left Shoal



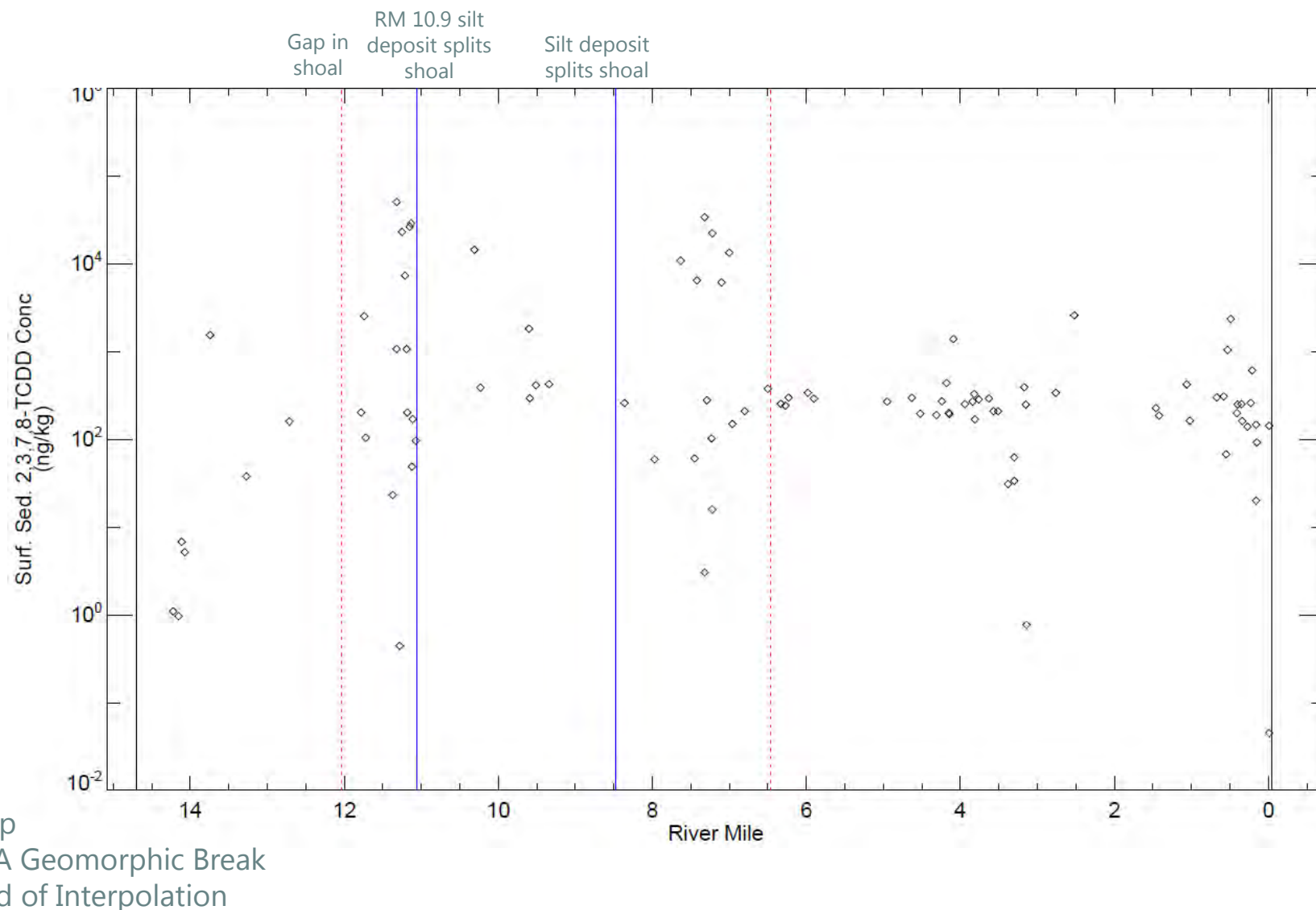
# River Segmentation – Left Shoal

2,3,7,8-TCDD Conc. (ng/kg)

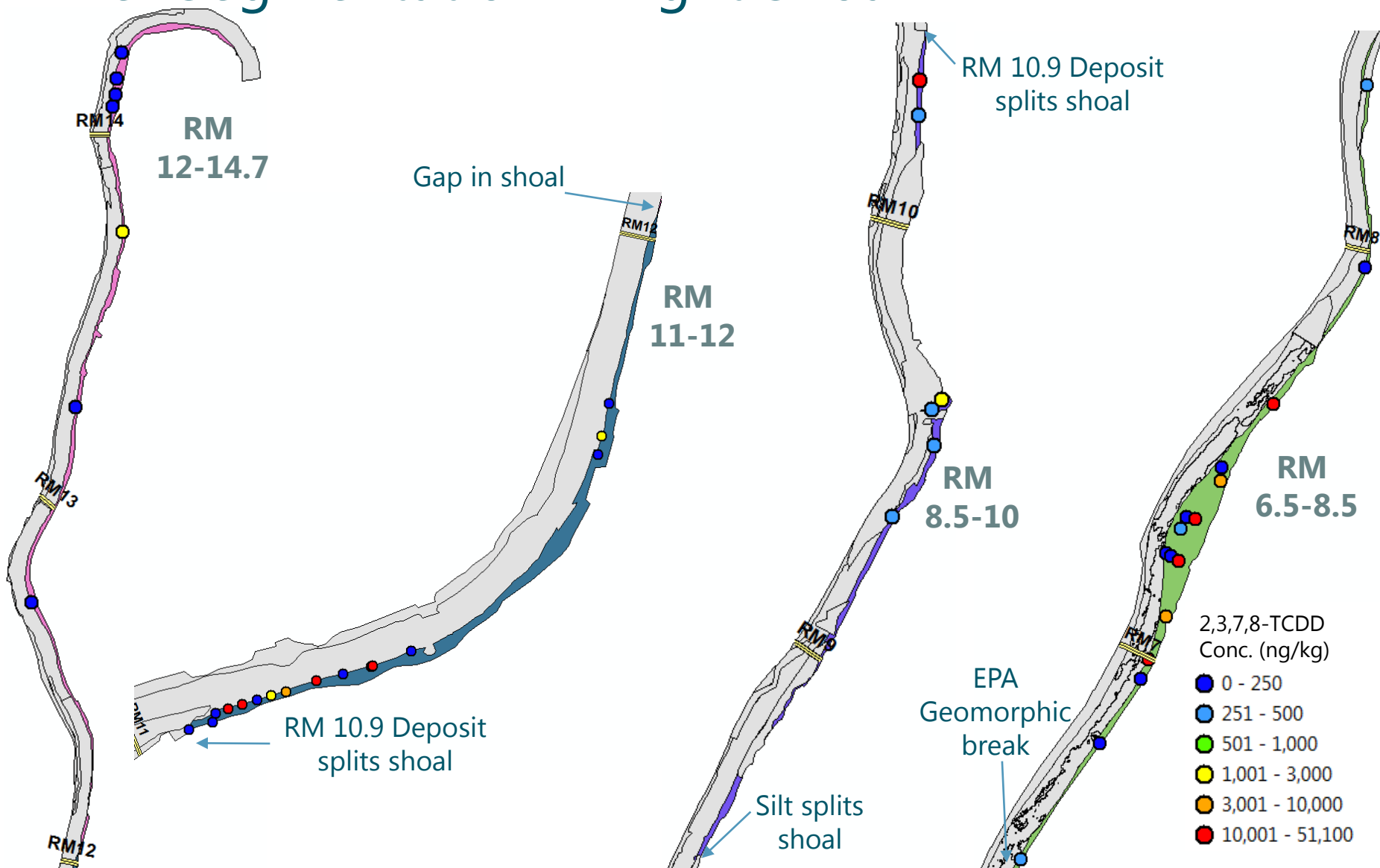
- 0 - 250
- 251 - 500
- 501 - 1,000
- 1,001 - 3,000
- 3,001 - 10,000
- 10,001 - 51,100



# River Segmentation – Right Shoal

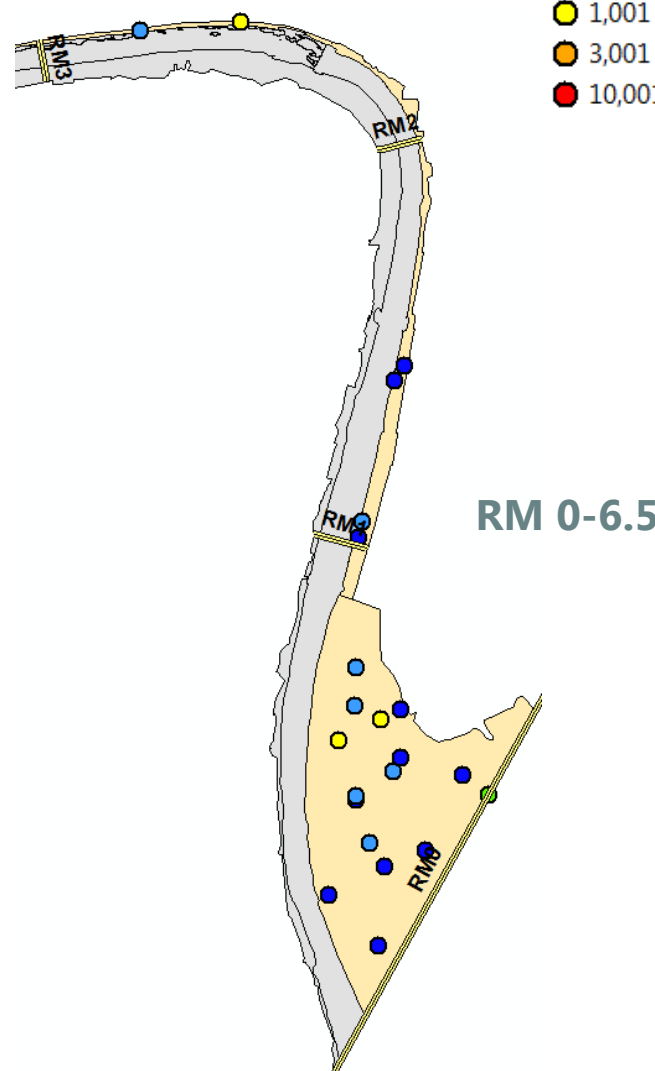
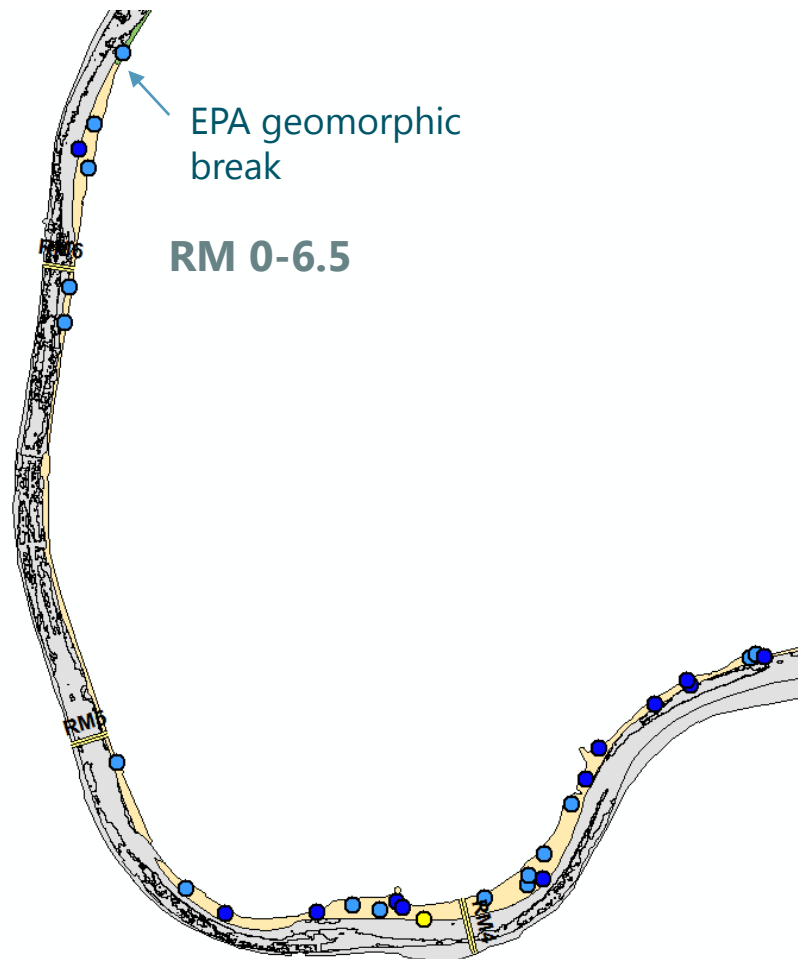
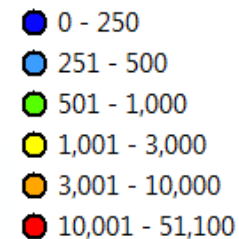


# River Segmentation – Right Shoal



# River Segmentation – Right Shoal

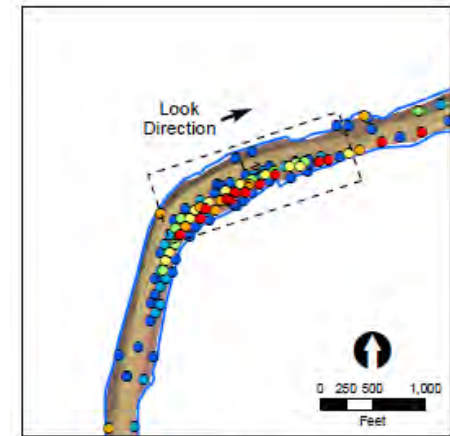
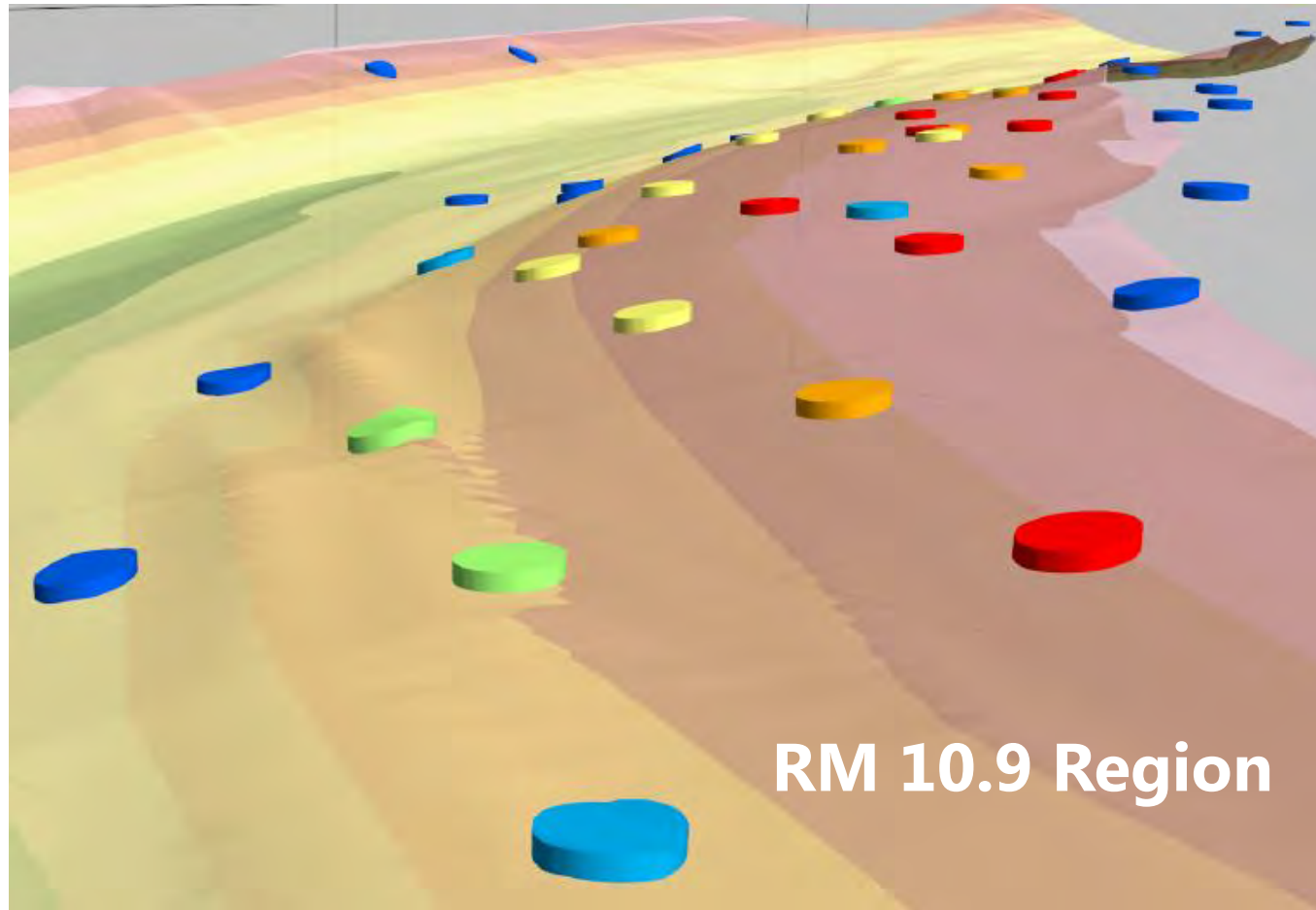
2,3,7,8-TCDD Conc. (ng/kg)



# Approach to developing a variogram

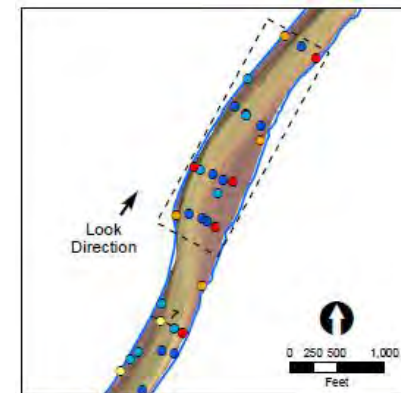
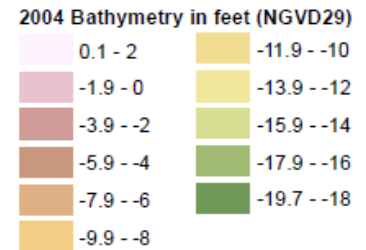
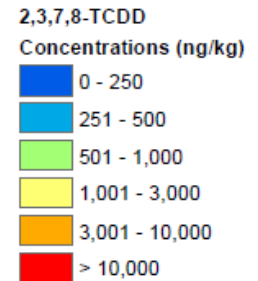
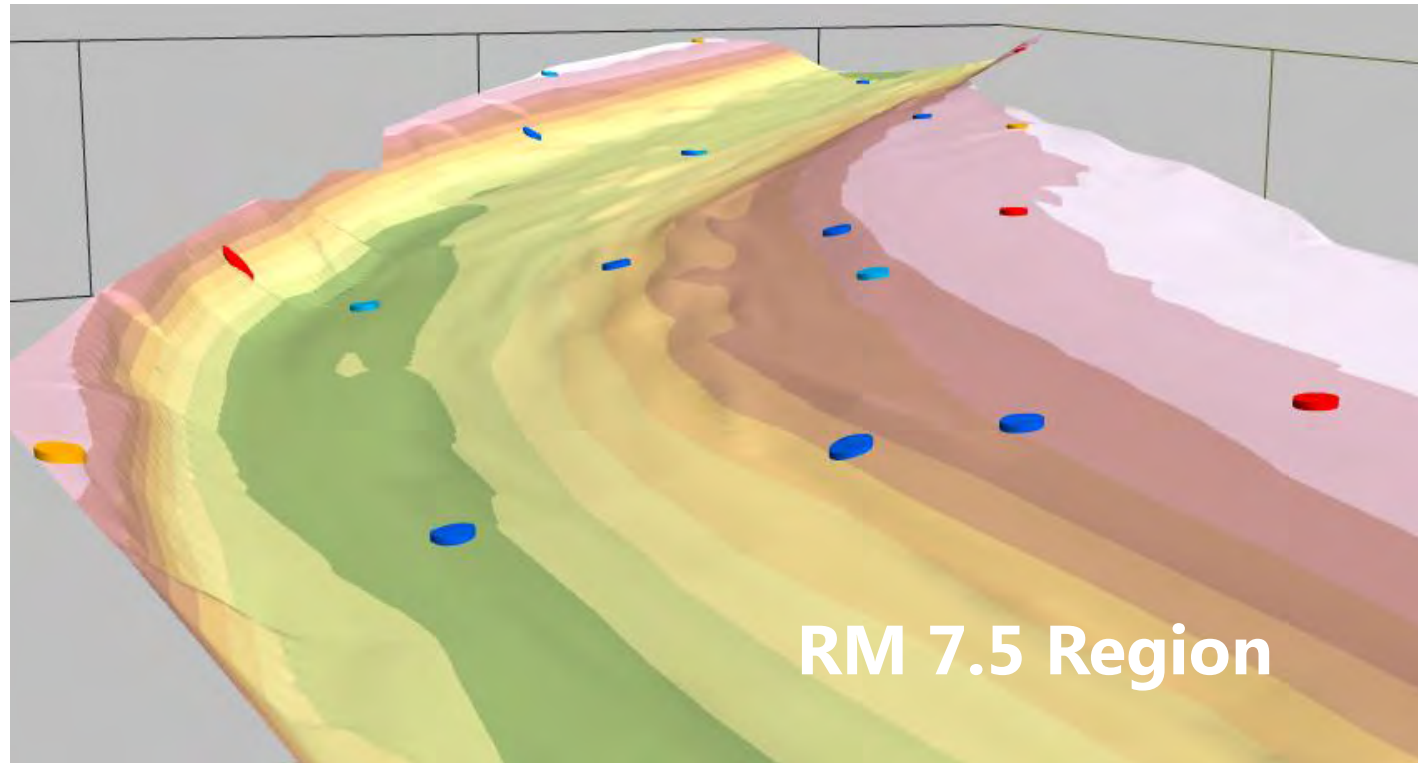
- Assess need for directional variogram
- Transform data to obtain approximate normal distribution
  - At present, using log transformation; considering benefit of using normal scores transformation
- “Straighten” the river via a coordinate transform
- Bin data by separation distance and calculate semi-variance in each bin
- Model the relationship of semivariance and separation distance

# Spatial Correlation is Anisotropic – Greater Along Flow than Across Flow



Vertical Exaggeration: 3X

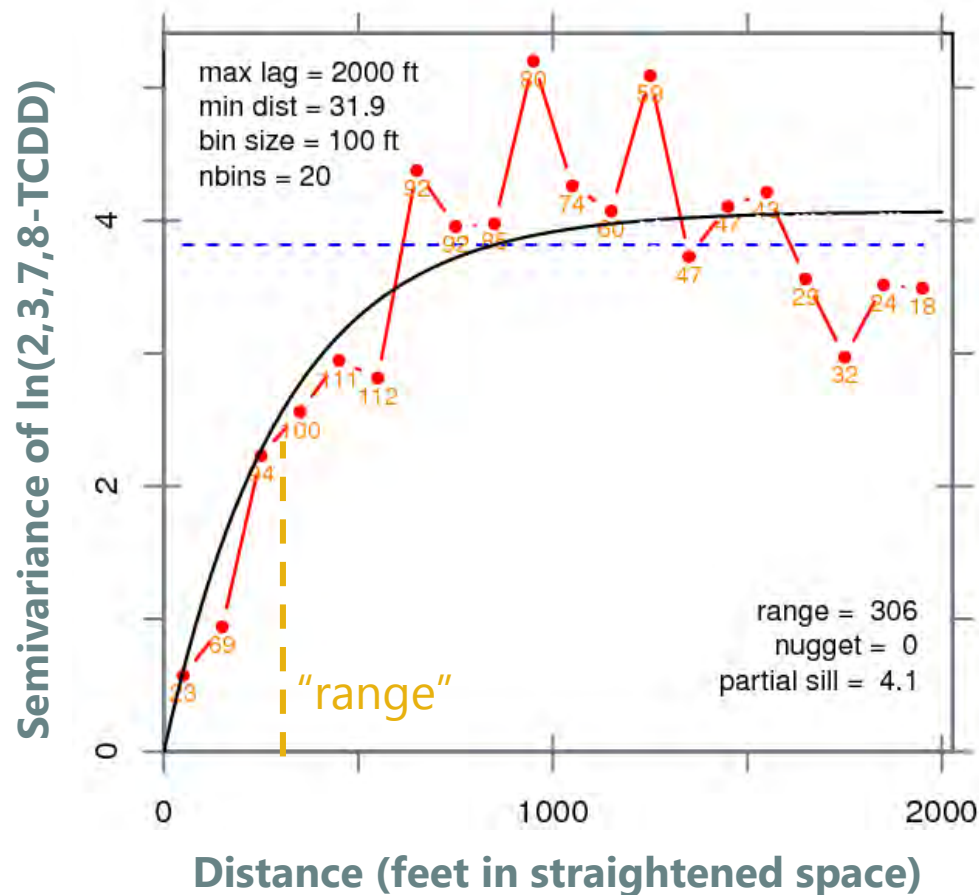
# Spatial Correlation is Anisotropic – Greater Along Flow than Across Flow



Vertical Exaggeration: 3X

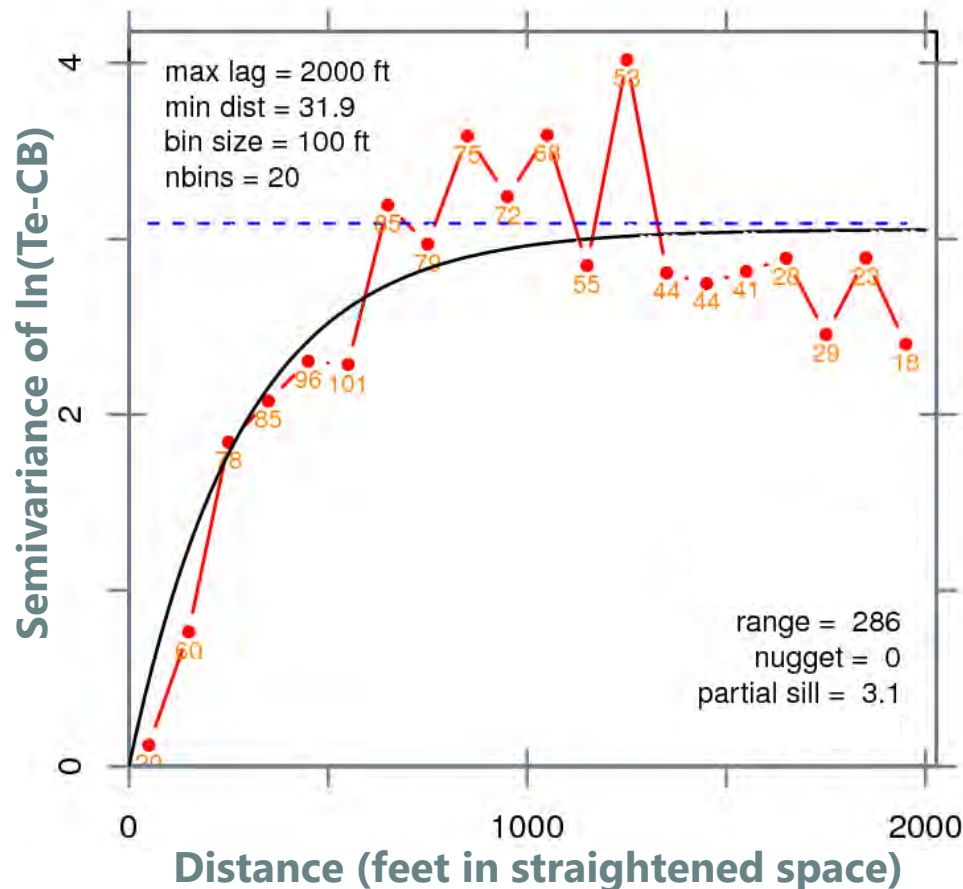
Approach used is to calculate along-flow variograms and assume anisotropy ratio to get cross-flow variograms. Ratio of 5 is used in work presented here.

# Along-Flow 2,3,7,8-TCDD Variogram at RM 10.9



Range defined here as distance to 63% of sill (per GeoR convention)

# Along-Flow Tetra-PCB Variogram at RM 10.9



Similarity to 2,3,7,8-TCDD variogram supports understanding of spatial correlation

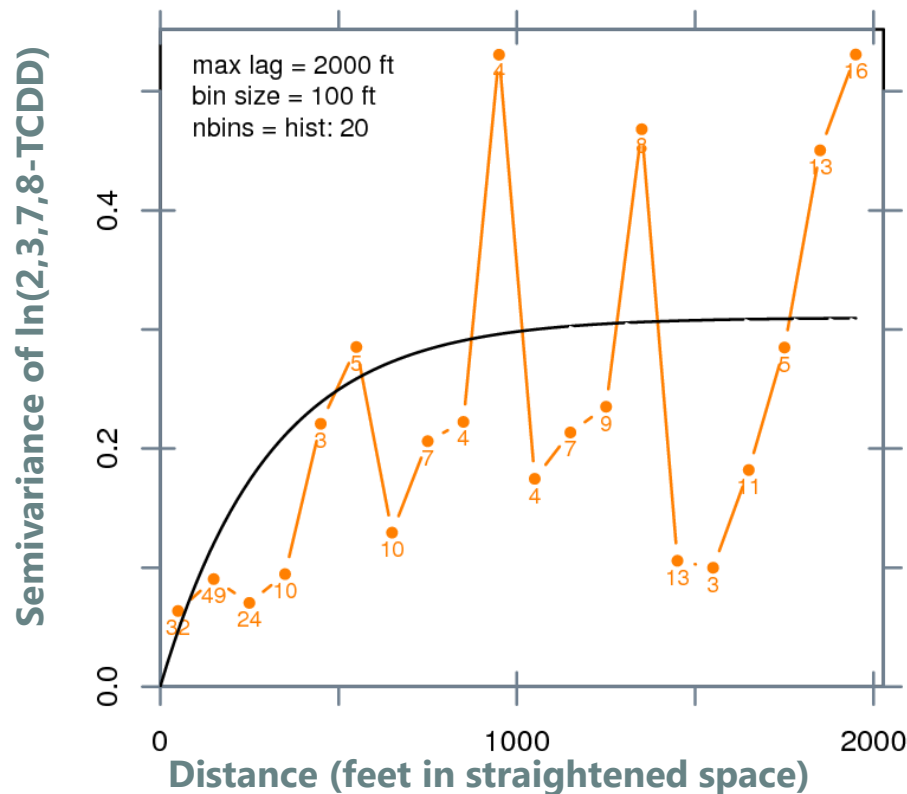
# Variogram Model for Other Areas

- Insufficient data to develop individual variograms for other areas
- Assume same shape as RM 10.9, but local variance
  - Note: In results shown here, local variance reduced in two groups to eliminate excessive influence of data at tails of distribution (Right Shoal RM 0-6.5 and Channel RM 13.75-14.7)

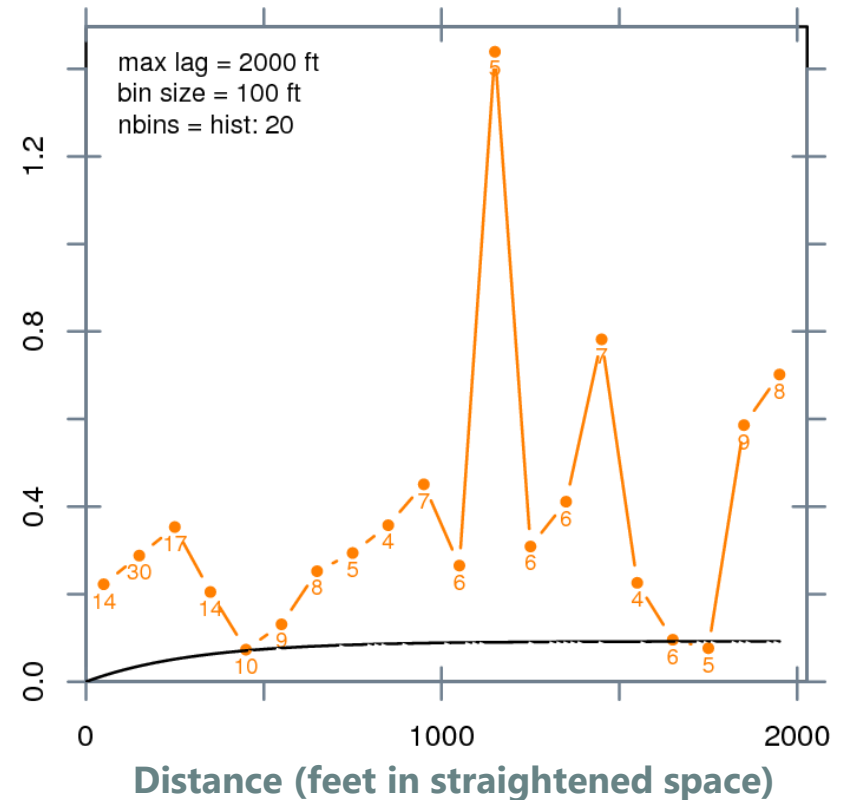
# Historical Data Support Applying RM 10.9 Variogram Shape to Other Areas

1995-2000 Data

Left Shoal RM 0-6

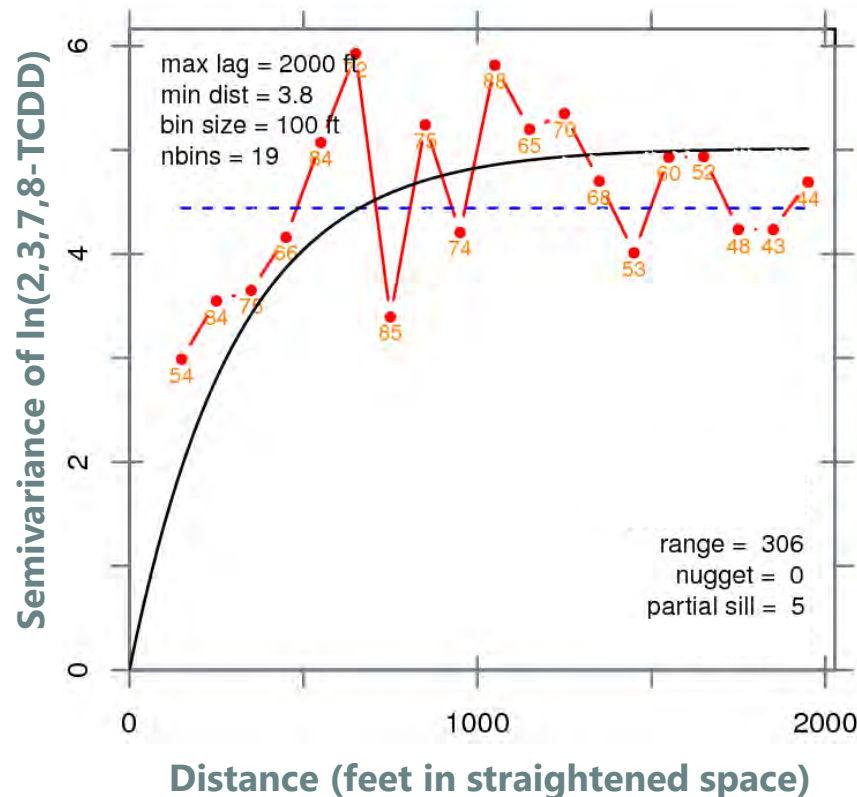


Right Shoal RM 0-6.5



# Phase 1 and Phase 2 Tierra Data Support Applying RM 10.9 Variogram Shape to Other Areas

## Phase 1&2 Data (0-2 ft segment) Tolerance = +/- 10 degrees



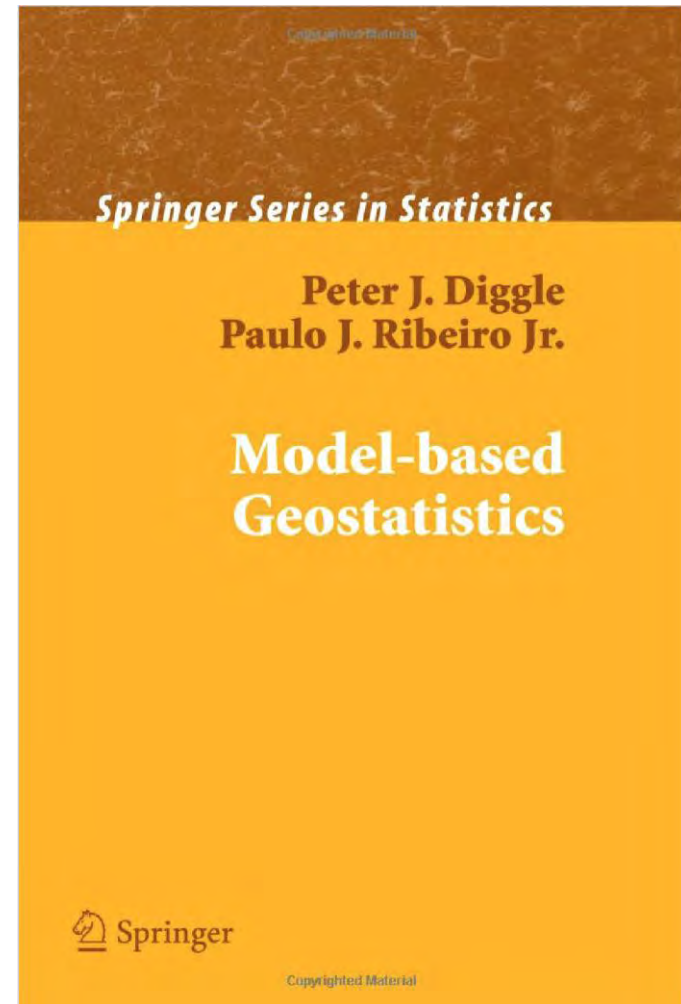
Comparison given less weight because of differing depth intervals and influence of ND data

# Kriging Approach

- Simple Kriging in log space
  - Trends removed by group delineation
  - Simple Kriging was chosen over Ordinary Kriging to reduce complications with lagrange multipliers and conditional simulation
  - Simple vs Ordinary Kriging predictions were compared and were very similar

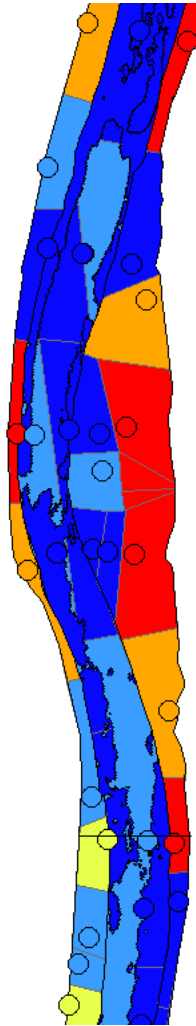
# Conditional Simulation Software

- Two Choices in R platform
  - GeoR
    - Bayesian Approach
  - Gstat
    - Sequential Gaussian Simulation
- GeoR was chosen
  - Used for variogram analysis
  - More computationally efficient
  - Book supporting its use

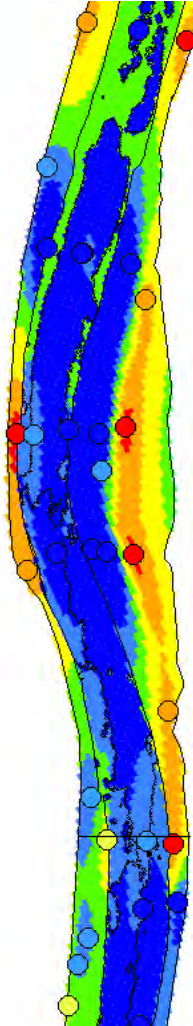


# Preliminary Results— Map RM 7.5

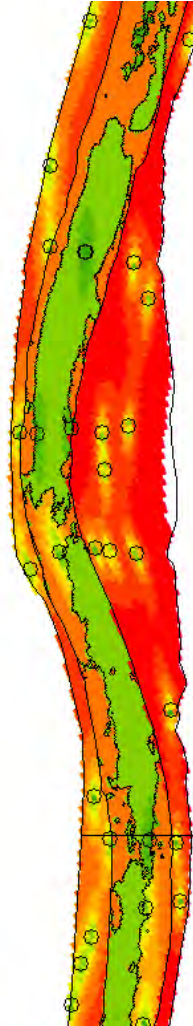
Thiessen Mapping



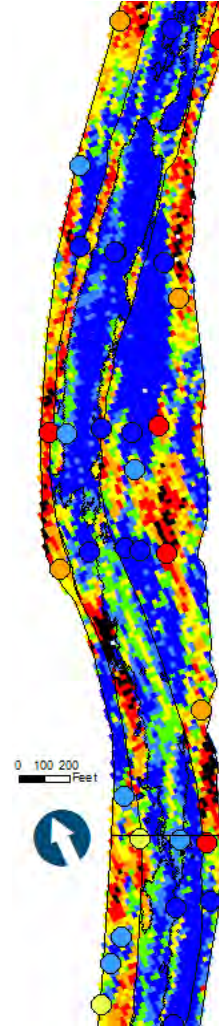
Kriging Median



Kriging Variance



Conditional Simulation 1



● Sample Location

— River Miles

□ Interpolation Boundaries

**2,3,7,8-TCDD Conc. (ng/kg)**

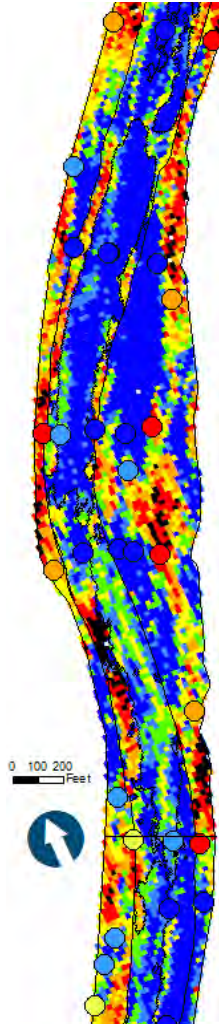
- 0 - 250
- 251 - 500
- 501 - 1,000
- 1,001 - 3,000
- 3,001 - 10,000
- 10,001 - 51,100
- 51,101 - 171,102

**Kriging Variance (Natural Log)**

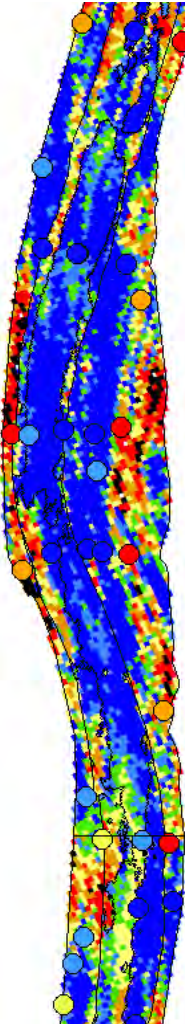
- 0.0 - 0.4
- 0.5 - 0.8
- 0.9 - 1.2
- 1.3 - 1.7
- 1.8 - 2.3
- 2.4 - 2.9
- 3.0 - 3.7
- 3.8 - 4.7
- 4.8 - 6.3
- 6.4 - 8.3

# Preliminary Results– Conditional Simulations

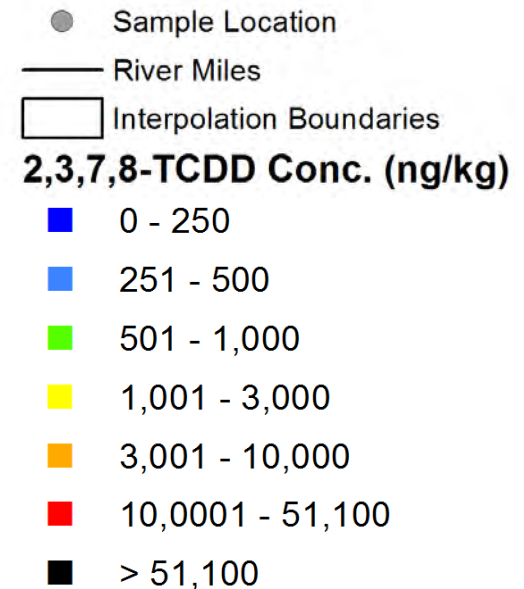
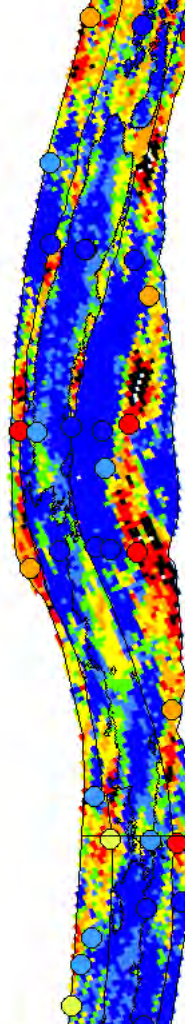
Conditional  
Simulation 1



Conditional  
Simulation 2

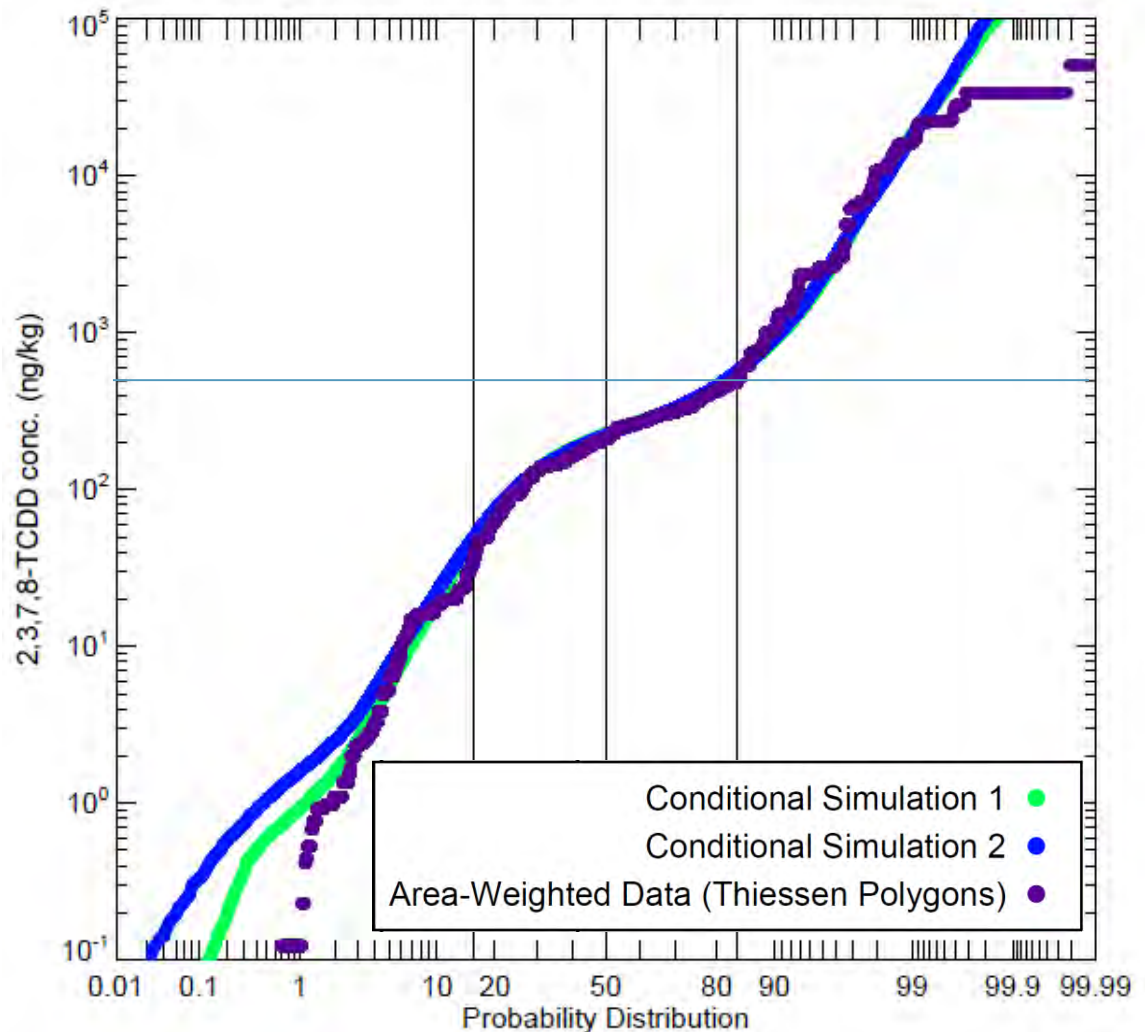


Conditional  
Simulation 3



# QC of Results - Concentration Distributions

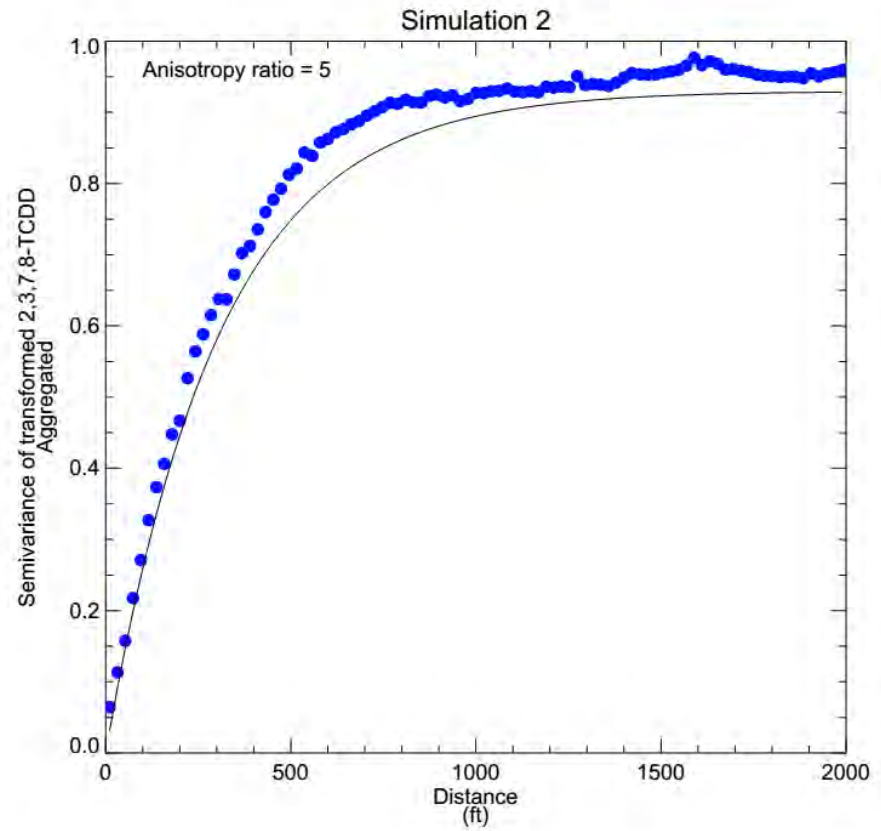
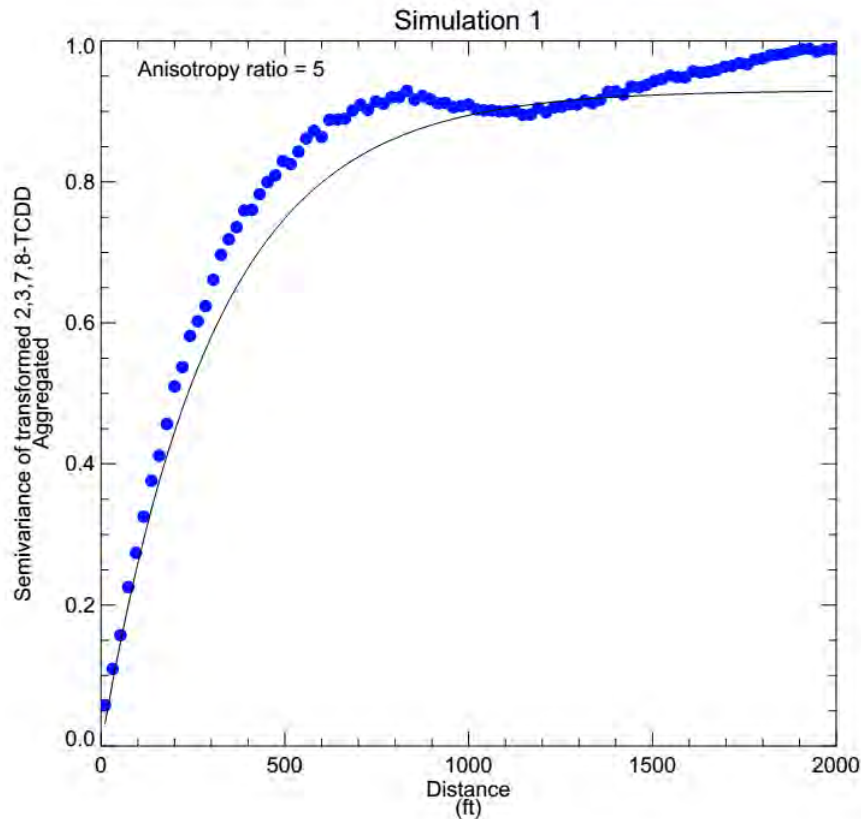
- CS recovers the concentration distributions
- Comparison to Area-Weighted Data



# QC of Results – Aggregate Variogram

- CS recovers the variogram

Note: Does not include Right Shoal RM 0-6.5 and Channel RM 13.75-14.7

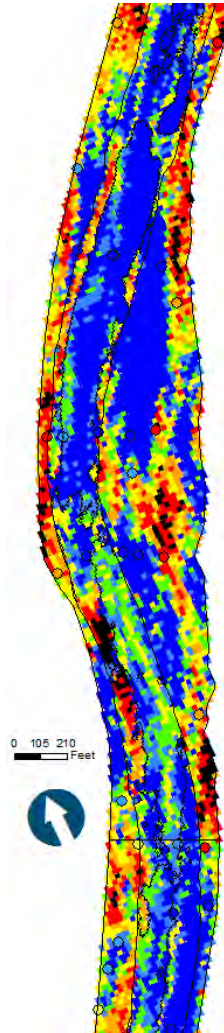


# Treatment of Simulation Results for Crafting a Targeted Remedy Alternative

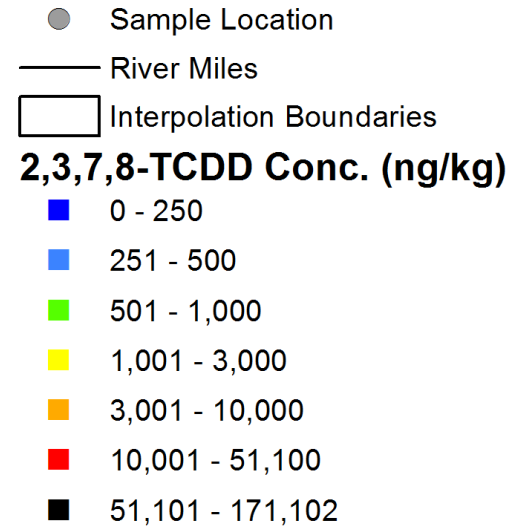
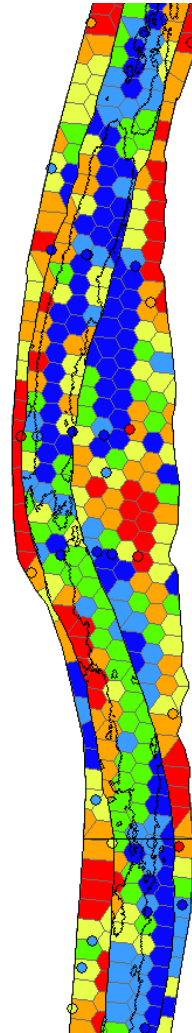
- Average results at 80-ft scale
  - Used as estimate of smallest remedial unit
- Cap concentrations at max. observed (51,100 ng/kg)
  - Occasional prediction of unrealistically high concentrations biases estimate of benefit achieved by targeted remediation

# Example of Proposed FS Approach (RM 7.5)

Conditional  
Simulation 1

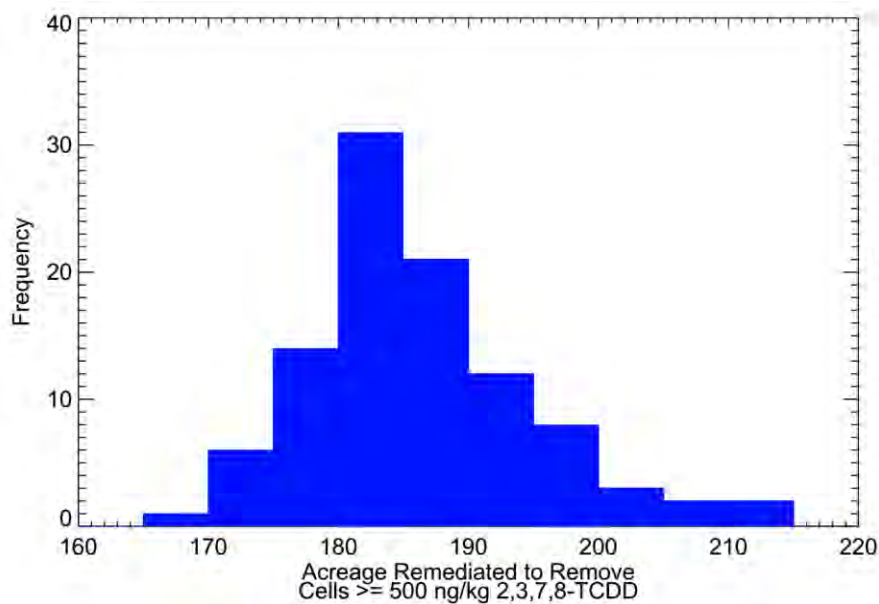


Conditional Simulation 1  
Averaged on 80-ft grid decision units

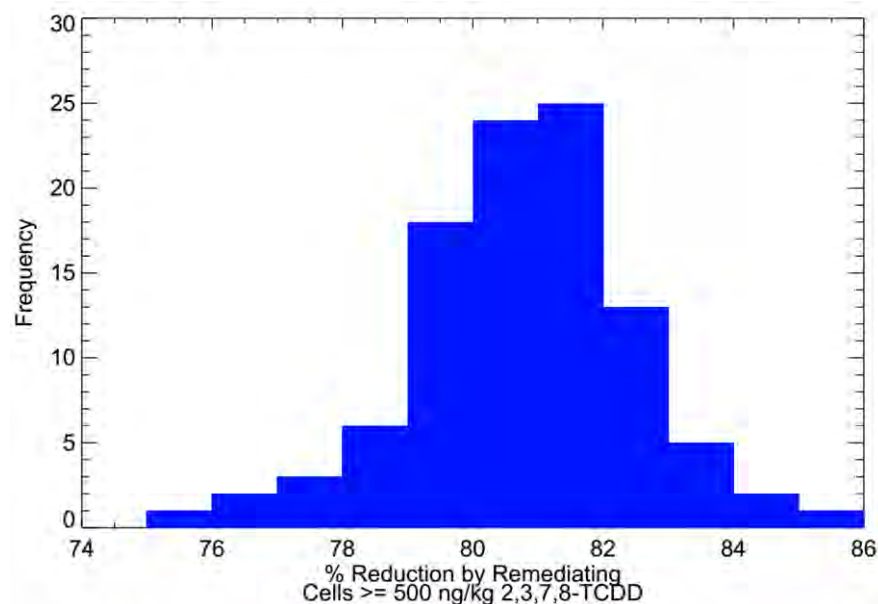


# Histograms – RAL 500 ng/kg

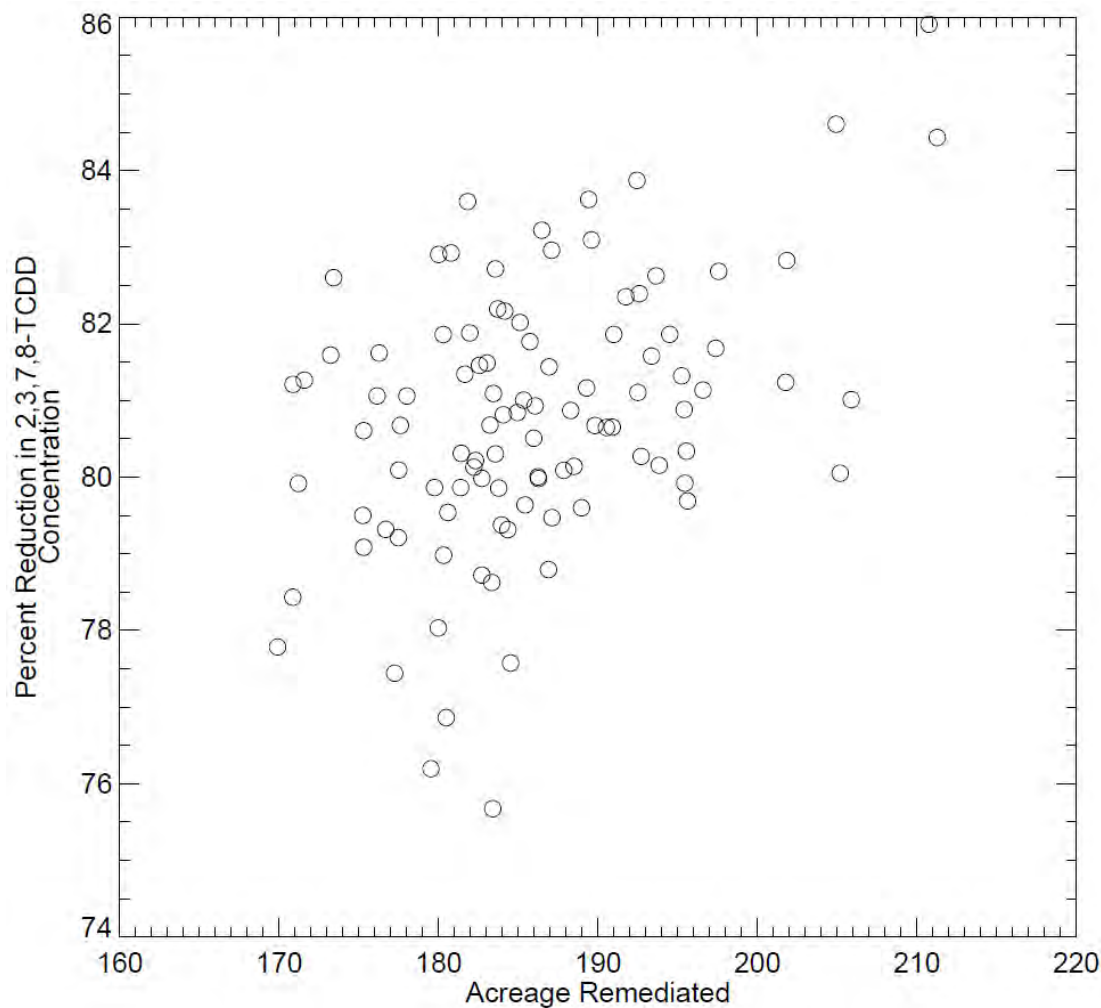
**Acres  
Remediated**



**% Reduction**



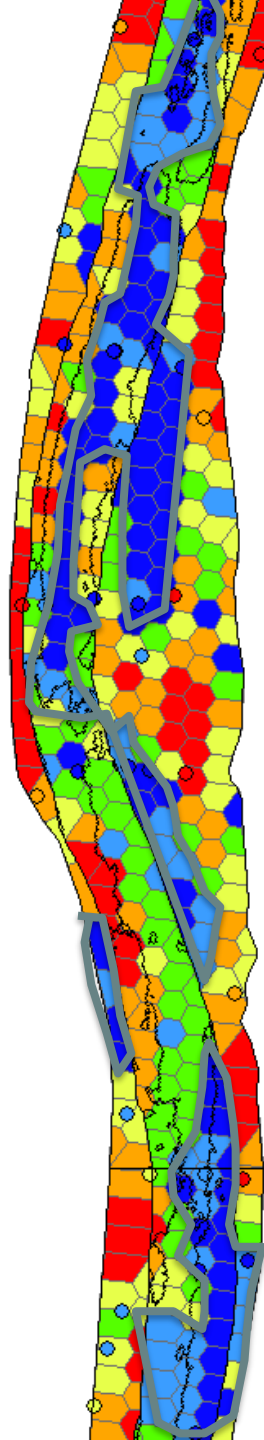
# Percent Reduction vs Acreage: 500 ng/kg RAL



# Summary

- Conditional simulation provides a means to quantify mapping uncertainty
- It provides information that can be used to make informed decisions that account for uncertainty
  - Choosing an RAL
  - Choosing areas meeting an RAL
  - Crafting a design sampling program aimed at efficiently reducing uncertainty
- Mapping using the LPR RI data set provides understanding sufficient to craft remedial alternatives for an FS
  - Uncertainty is reasonable and can be reduced during remedial design

Illustration of  
Delineation of  
Remedial  
Footprint for an  
FS Alternative



# Backup Slides

# Effect of Trimming Tails of the Right Shoal RM 0-6.5 Sample Data on the Variogram and its Comparison to Historical Data

